

PATENT

Attorney Docket No. UCAL-246/02US

1 NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS  
2  
3

4 ACKNOWLEDGMENTS

5 This invention was supported in part by grants from the National Institutes of Health  
6 grant number 1 R01 DK 43787 and 5 R01 DK 41842. The U.S. Government may have  
7 certain rights in this invention.

8 CROSS-REFERENCE TO RELATED APPLICATIONS  
9

10 This application claims the benefit of the following provisional applications: United  
11 States Ser. No. 60/008,540 and 60/008,543, filed December 13, 1995, and Ser. No.  
12 60/008,606, filed December 14, 1995. This application claims the benefit of the following  
13 U.S. patent application: United States Ser. No. 08/764,870, filed December 13, 1996.  
14

15 INTRODUCTION

16 Technical Field

17 This invention relates to computational methods for designing ligands that bind to  
18 nuclear receptors, crystals of nuclear receptors, synthetic ligands of nuclear receptors and  
19 methods of using synthetic ligands.

20 Background

21 Nuclear receptors represent a superfamily of proteins that specifically bind a  
22 physiologically relevant small molecule, such as hormone or vitamin. As a result of a  
23 molecule binding to a nuclear receptor, the nuclear receptor changes the ability of a cell to  
24 transcribe DNA, i.e. nuclear receptors modulate the transcription of DNA, although they  
25 may have transcription independent actions. Unlike integral membrane receptors and  
26 membrane associated receptors, the nuclear receptors reside in either the cytoplasm or  
27 nucleus of eukaryotic cells. Thus, nuclear receptors comprise a class of intracellular, soluble  
28 ligand-regulated transcription factors.

1 Nuclear receptors include receptors for glucocorticoids (GRs), androgens (ARs),  
2 mineralocorticoids (MRs), progestins (PRs), estrogens (ERs), thyroid hormones (TRs),  
3 vitamin D (VDRs), retinoids (RARs and RXRs), peroxisomes (XPARs and PPARs) and  
4 icosanoids (IRs). The so called "orphan receptors" are also part of the nuclear receptor  
5 superfamily, as they are structurally homologous to the classic nuclear receptors, such as  
6 steroid and thyroid receptors. To date, ligands have not been identified with orphan  
7 receptors but it is likely that small molecule ligands will be discovered in the near future for  
8 this class of transcription factors. Generally, nuclear receptors specifically bind  
9 physiologically relevant small molecules with high affinity and apparent  $K_d$ 's are commonly  
10 in the 0.01 - 20 nM range, depending on the nuclear receptor/ligand pair.

11 Development of synthetic ligands that specifically bind to nuclear receptors has been  
12 largely guided by the trial and error method of drug design despite the importance of nuclear  
13 receptors in a myriad of physiological processes and medical conditions such as hypertension,  
14 inflammation, hormone dependent cancers (e.g. breast and prostate cancer), modulation of  
15 reproductive organ function, hyperthyroidism, hypercholesterolemia and obesity. Previously,  
16 new ligands specific for nuclear receptors were discovered in the absence of information on  
17 the three dimensional structure of a nuclear receptor with a bound ligand. Before the present  
18 invention, researchers were essentially discovering nuclear receptor ligands by probing in the  
19 dark and without the ability to visualize how the amino acids of a nuclear receptor held a  
20 ligand in its grasp.

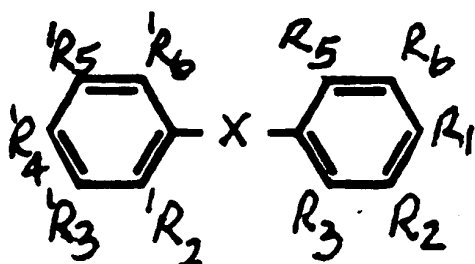
21 Consequently, it would be advantageous to devise methods and compositions for  
22 reducing the time required to discover ligands to nuclear receptors, synthesize such  
23 compounds and administer such compounds to organisms to modulate physiological processes  
24 regulated by nuclear receptors.  
25  
26

1 SUMMARY OF THE INVENTION

2 The present invention provides for crystals of nuclear receptor ligand binding domains  
3 with a ligand bound to the ligand binding domain (LBD). The crystals of the present  
4 invention provide excellent atomic resolution of the amino acids that interact with nuclear  
5 receptor ligand, especially thyroid receptor ligands. The three dimensional model of a  
6 nuclear receptor LBD with a ligand bound reveals a previously unknown structure for nuclear  
7 receptors and shows that the ligand is bound in a water inaccessible binding cavity of the  
8 ligand binding domain of the nuclear receptor.

9 The present invention also provides for computational methods using three  
10 dimensional models of nuclear receptors that are based on crystals of nuclear receptor LBDs.  
11 Generally, the computational method of designing a nuclear receptor ligand determines which  
12 amino acid or amino acids of a nuclear receptor LBD interact with a chemical moiety (at  
13 least one) of the ligand using a three dimensional model of a crystallized protein comprising  
14 a nuclear receptor LBD with a bound ligand, and selecting a chemical modification (at least  
15 one) of the chemical moiety to produce a second chemical moiety with a structure that either  
16 decreases or increases an interaction between the interacting amino acid and the second  
17 chemical moiety compared to the interaction between the interacting amino acid and the  
18 corresponding chemical moiety on the natural hormone.

19 Also provided is a method of modulating the activity of a nuclear receptor. The  
20 method can be *in vitro* or *in vivo*. The method comprises administering *in vitro* or *in vivo* a  
21 sufficient amount of a compound of the following formula:



29 FORMULA I.

30 where the compound fits specially and preferentially into a nuclear hormone receptor LBD of  
31 interest. The method is exemplified by modulating the activity of a thyroid receptor (TR).

1 For modulating TR activity, a compound of Formula I is employed that fits spacially and  
2 preferentially into a TR ligand binding domain (TR LBD), including compounds specific for  
3 a TR LBD isoform of interest. Of particular interest are the TR LBD isoforms  $\alpha$  (TR- $\alpha$ ) and  
4  $\beta$  (TR- $\beta$ ). Additional compounds of interest include derivatives of Formula I, such as those  
5 compounds having the biphenyl ( $\phi$ -X- $\phi$ ) or single phenyl ( $\phi$ -X or X- $\phi$ ) nucleus of Formula I  
6 and its corresponding substituent groups described herein. Compounds that are iteratively  
7 designed using structural information gleaned from these compounds and which modulate  
8 nuclear hormone receptor activity also are of interest.

9 The present invention also includes a method for identifying a compound capable of  
10 selectively modulating the activity of a nuclear receptor. This aspect of the invention is  
11 exemplified by a method for identifying a compound capable of selectively modulating the  
12 activity of a TR isoform. The method comprises modeling test compounds that fit spacially  
13 and preferentially into a TR LBD isoform of interest using an atomic structural model of a  
14 TR LBD isoform bound to a test compound, screening the test compounds in a biological  
15 assay for TR isoform activity characterized by binding of a test compound to a TR LBD  
16 isoform, and identifying a test compound that selectively modulates the activity of a TR  
17 isoform. The compounds may be those of Formula I or derivatives thereof, including  
18 compounds having a biphenyl or single phenyl nucleus of Formula I.

19 Further included is a method for identifying agonist or antagonist ligands of a nuclear  
20 receptor using the atomic coordinates of a LBD in conjunction with a computerized modeling  
21 system. This aspect of the invention is exemplified by identifying a TR agonist or antagonist  
22 ligand by providing the atomic coordinates of a TR LBD to a computerized modeling system,  
23 modeling ligands which fit spacially into the TR LBD, and identifying in a biological assay  
24 for TR activity a ligand which increases or decreases TR activity. The compounds can be  
25 those of Formula I or derivatives thereof, including compounds having a biphenyl or single  
26 phenyl nucleus of Formula I.

27 Also provided is a method of identifying a compound that selectively modulates the  
28 activity of one type of nuclear receptor compared to other nuclear hormone receptors. The  
29 method is exemplified by modeling test compounds which fit spacially into a TR LBD using  
30 an atomic structural model of a TR LBD, selecting a compound comprising conformationally  
31 constrained structural features that interact with conformationally constrained residues of a

1 TR LBD, and identifying in a biological assay for TR activity a compound that selectively  
2 binds to a TR LBD compared to other nuclear receptors. The conformationally constrained  
3 features involved in receptor-selective ligand binding can be identified by comparing atomic  
4 models of receptor isoforms bound to the same and/or different ligands. The methods  
5 facilitate design and selection of compounds that have increased selectivity for a particular  
6 nuclear receptor. The compounds may be those of Formula I or derivatives thereof,  
7 including compounds having the biphenyl or single phenyl nucleus of Formula I.

8 Another aspect of the invention is a method for increasing the receptor selectivity of a  
9 compound for a particular type of nuclear receptor. This involves the chemical modification  
10 of a substituent group of a compound of Formula I to generate compounds which have  
11 increased selectivity for one type of receptor. For example, chemical modification of a  
12 substituent group of the compound of Formula I can be used to introduce additional  
13 constraints into a compound that modulates TR activity to increase its selectivity *in vivo* for  
14 TR-type receptors. Additional constraints also may be added for stability. The modified  
15 groups will preferably interact with a conformationally constrained structural feature of a TR  
16 LBD that is conserved among TR isoforms. A more preferred method comprises selecting  
17 compounds having conformationally constrained groups that interact with conformationally  
18 constrained residues of a TR LBD conserved among TR isoforms. The compounds can be  
19 those of Formula I or derivatives thereof, including compounds having the biphenyl or single  
20 phenyl nucleus of Formula I.

21 The invention finds use in the selection and characterization of peptide,  
22 peptidomimetic or synthetic compounds identified by the methods of the invention,  
23 particularly new lead compounds useful in treating disorders related to nuclear receptor-based  
24 deficiencies, including TR-related disorders. For TR-related disorders, the compounds and  
25 methods of the invention can be used to modulate TR activity by administering to a mammal  
26 in need thereof a sufficient amount of compound of Formula I or derivative thereof that fits  
27 spacially and preferentially into a TR LBD.

#### 29 BRIEF DESCRIPTION OF THE DRAWINGS

30 FIG. 1 is a diagram illustrating computational methods for designing ligands that  
31 interact with nuclear receptors of the nuclear receptor superfamily.

1 **FIG. 2** is a schematic representation of nuclear receptor structures, indicating regions  
2 of homology within family members and functions of the various domains.

3 ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup> shows the aligned amino acid sequences of the ligand binding domains of  
4 several members of the nuclear receptor superfamily. <sup>23</sup>

5 **FIG. 4** is a ribbon drawing of the rat TR- $\alpha$  LBD with secondary structure elements  
6 labelled. The ligand (magenta) is depicted as a space-filling model. Alpha helices and coil  
7 conformations are yellow, beta strands are blue.

8 **FIG. 5** shows two cross-sections of a space-filling model of rat TR- $\alpha$  exposing the  
9 ligand (magenta) tightly packed within the receptor.

10 **FIG. 6** is a schematic of the ligand binding cavity. Residues which interact with the  
11 ligand appear approximately at the site of interaction. Hydrogen bonds are shown as dashed  
12 lines between the bonding partners; distances for each bond are listed. Non-bonded contacts  
13 are shown as radial spokes which face toward interacting atoms.

14 **FIG. 7** is the distribution of crystallographic temperature factors in the refined rat  
15 TR- $\alpha$  LBD. The distribution is represented as a color gradation ranging from less than 15  
16 (dark blue) to greater than 35 (yellow-green).

17 **FIG. 8** is a ribbon drawing of the rat TR- $\alpha$  LBD showing the c-terminal activation  
18 domain to ligand. Residues which comprise the c-terminal activation domain (Pro393-  
19 Phe405) are depicted as a stick representation. Hydrophobic residues, particularly Phe401  
20 and Phe405 (blue) face inwards toward the ligand. Glu403 (red) projects outward into the  
21 solvent.

22 **FIG. 9** is an electrostatic potential surface of the rat TR- $\alpha$  LBD, calculated using  
23 GRAPH. Negative electrostatic potential is red; positive electrostatic potential is blue. The  
24 c-terminal activation domain forms a largely hydrophobic (white). The Glu403 is presented  
25 as a singular patch of negative charge (red).

26 **FIG. 10** is a diagram comparing agonists and antagonists for several nuclear  
27 receptors.

28 **FIG. 11** is the synthetic scheme for preparation of TS1, TS2, TS3, TS4 and TS5.

29 **FIG. 12** is the synthetic scheme for preparation of TS6 and TS7.

30 **FIG. 13** is the synthetic scheme for preparation of TS8.

31 ~~FIG. 14~~ <sup>FIGS. 14A-14B</sup> is the synthetic scheme for preparation of TS10.

1        **FIG. 15** depicts the chemical structures of several TR ligands.

2        **FIG. 16** is a graph illustrating competition assays in which T<sub>3</sub> and Triac compete with  
3 labeled T<sub>3</sub> for binding to human TR- $\alpha$  or human TR- $\beta$ .

4        **FIG. 17** depicts a Scatchard analysis of labelled T<sub>3</sub> binding to TR- $\alpha$  and TR- $\beta$ .

5        **FIG. 18** is a chart showing the effect of TS-10 on the transcriptional regulation of the  
6 DR4-ALP reporter gene in the presence or absence of T3 as assayed in TRAF $\alpha$ 1 reporter  
7 cells.

8        **FIG. 19** is a chart showing the effect of TS-10 on the transcriptional regulation of the  
9 DR4-ALP reporter gene in the presence or absence of T3 as assayed in TRAF $\beta$ 1 reporter  
10 cells.

11        **FIG. 20** is a chart showing the effect of TS-10 on the transcriptional regulation of the  
12 DR4-ALP reporter gene in the presence or absence of T3 as assayed in HepG2, a liver  
13 reporter cell line.

14        **FIG. 21** is a partial ribbon drawing of TR- $\alpha$  LBD with T3 in the ligand binding  
15 cavity. Selected interacting amino acids are labelled, including Ile221, Ile222 and Ser260,  
16 Ala263, Ile299 and Leu 276.

17        **FIG. 22** is a partial ribbon drawing of TR- $\alpha$  LBD with T3 and Dimit superimposed in  
18 the ligand binding cavity. Interactions with Ile221, Ile222, Ala260, Ile 299 and Leu276 are  
19 labelled.

20        **FIG. 23** is a partial ribbon drawing of TR- $\alpha$  LBD with T3, illustrating the three  
21 Arginine residues (Arg228, Arg262 and Arg 266 (dark stick figures)) of the polar pocket,  
22 three water molecules HOH502, HOH503 and HOH504, with hydrogen bonds indicated by  
23 dotted lines.

24        **FIG. 24** is a partial ribbon drawing of TR- $\alpha$  LBD with Triac, illustrating the three  
25 Arginine residues (dark stick figures) of the polar pocket, water molecules (HOH503,  
26 HOH504 and HOH600), with hydrogen bonds indicated by dotted lines.

27        **FIG. 25** is a partial ribbon drawing of the TR- $\alpha$  LBD with T3 and Triac  
28 superimposed in the ligand binding cavity. The drawing shows several interacting amino  
29 acid residues in the polar pocket that remain unchanged whether T3 or Triac occupies the  
30 ligand binding cavity: Arg262, Asn179, HOH503 and HOH504, and Ser277. Both Arg228  
31 and Arg 266 occupy two different positions, depending on whether T3 or Triac is bound.

1        **FIG. 26A and 26B** are stereochemical representations of the TR- $\alpha$  LBD with Dimit  
2 bound.

3        **FIG. 27** is a partial ribbon drawing of TR- $\beta$  LBD with GC-1 in the ligand binding  
4 cavity. Amino acids Arg282, Arg316, Arg320, Asn 331 and His435 are labelled.

5        **FIG. 28** is a partial ribbon drawing of TR- $\beta$  LBD with Triac in the ligand binding  
6 cavity. Amino acids Arg282, Arg316, Arg320, Asn331 and His435 are labelled.

7        **FIG. 29** is a partial ribbon drawing of TR- $\beta$ LBD with GC-1 (Blue) overlayed with  
8 TR- $\alpha$  LBD with Dimit (Red) in the ligand binding cavities. Amino acids Arg228, Arg262,  
9 Arg266 and Ser277 (TR- $\alpha$  LBD), and Arg282, Arg316, Arg320 and Asn331 (TR- $\beta$  LBD)  
10 are labelled.

11        **FIG. 30** is a partial ribbon drawing of TR- $\beta$  LBD with Triac (Blue) overlayed with  
12 TR- $\alpha$  LBD with Triac (Red) in the ligand binding cavities. Amino acids Arg228, Arg262,  
13 Arg266, Ser277 and His381 (TR- $\alpha$  LBD), and Arg282, Arg316, Arg320 and His435 (TR- $\beta$   
14 LBD) are labelled.

15        **FIG. 31** is a graph showing competition curves comparing wildtype TR- $\alpha$  and TR- $\beta$   
16 to a variant TR- $\beta$  having a single amino acid substitution in the ligand binding domain.

17        **FIG. 32** shows atomic numbering for thyronine-like ligands.

18        **APPENDIX 1** is an appendix of references.

19        **APPENDIX 2** is a chart of amino acids that interact with a TR ligand, for TR  
20 complexed with Dimit, Triac, IpBr<sub>2</sub>, T<sub>3</sub> and GC-1.

21        **APPENDIX 3** is a chart of atomic coordinates for the crystal of rat TR- $\alpha$  LBD  
22 complexed with Dimit.

23        **APPENDIX 4** is a chart of atomic coordinates for the crystal of rat TR- $\alpha$  LBD  
24 complexed with Triac.

25        **APPENDIX 5** is a chart of atomic coordinates for the crystal of rat TR- $\alpha$  LBD  
26 complexed with IpBr<sub>2</sub>.

27        **APPENDIX 6** is a chart of atomic coordinates for the crystal of rat TR- $\alpha$  LBD  
28 complexed with T<sub>3</sub>.

29        **APPENDIX 7** is a chart of atomic coordinates for the crystal of human TR- $\beta$  LBD  
30 complexed with Triac.



1 APPENDIX 8 is a chart of atomic coordinates for the crystal of human TR- $\beta$ -LBD  
2 complexed with GC-1.

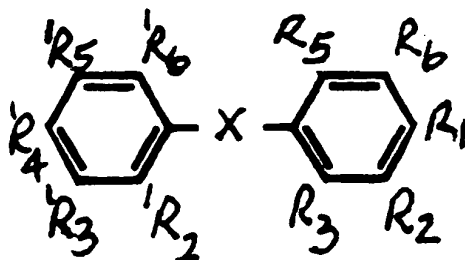
3  
4 DETAILED DESCRIPTION OF THE INVENTION

5 INTRODUCTION

6 The present invention provides new methods, particularly computational methods, and  
7 compositions for the generation of nuclear receptor synthetic ligands based on the three  
8 dimensional structure of nuclear receptors, particularly the thyroid receptor (herein referred  
9 to as "TR"). Previously, the lack of three dimensional structural information about the  
10 ligand binding domain of a nuclear receptor thwarted the field of nuclear receptor drug  
11 discovery, especially the absence of three dimensional structural information relating to a  
12 nuclear receptor with a ligand bound.

13 Described herein for the first time are crystals and three dimensional structural  
14 information from a nuclear receptor's ligand binding domain (LBD) with a ligand bound.  
15 The structure of the TR LBD complexed with 3,5,3'-triiodothyronine ( $T_3$ ), 3,5-dibromo-3'-  
16 isopropylthyronine (IpBr<sub>2</sub>), 3,5- dimethyl-3'-isopropylthyronine (Dimit), and 3,5,3'-  
17 triiodothyroacetic acid (Triac), 3,5-dimethyl-4-(4'-hydroxy-3'isopropylbenzyl)-phenoxy acetic  
18 acid (GC1) are exemplified. Such crystals offer superior resolution at the atomic level and  
19 the ability to visualize the coordination of nuclear receptor ligands by amino acids that  
20 comprise the LBD. The present invention also provides computational methods for designing  
21 nuclear receptor synthetic ligands using such crystal and three dimensional structural  
22 information to generate synthetic ligands that modulate the conformational changes of a  
23 nuclear receptor's LBD. Such synthetic ligands can be designed using the computational  
24 methods described herein and shown, in part, in FIG. 1. These computational methods are  
25 particularly useful in designing an antagonist or partial agonist to a nuclear receptor, wherein  
26 the antagonist or partial agonist has an extended moiety that prevents any one of a number of  
27 ligand-induced molecular events that alter the receptor's influence on the regulation of gene  
28 expression, such as preventing the normal coordination of the activation domain observed for  
29 a naturally occurring ligand or other ligands that mimic the naturally occurring ligand, such  
30 as an agonist. As described herein, synthetic ligands of nuclear receptors will be useful in  
31 modulating nuclear receptor activity in a variety of medical conditions.

Of particular interest is use of such ligands in a method of modulating TR activity in a mammal by administering to a mammal in need thereof a sufficient amount of a compound of Formula I,



where the compound fits spatially and preferentially into a TR LBD. By "fits spatially" is intended that the three-dimensional structure of a compound is accommodated geometrically by a cavity or pocket of a TR LBD. By "TR LBD" is intended a structural segment or segments of thyroid hormone receptor polypeptide chain folded in such a way so as to give the proper geometry and amino acid residue configuration for ligand binding. This is the physical arrangement of protein atoms in three-dimensional space forming a ligand binding pocket or cavity. By "fits spatially and preferentially" is intended that a compound possesses a three-dimensional structure and conformation for selectively interacting with a TR LBD. Compounds of interest also include derivatives of Formula I. By "derivatives of Formula I" is intended compounds that comprise at least a single phenyl scaffold ( $\phi$ -X or X- $\phi$ ) of the biphenyl scaffold ( $\phi$ -X- $\phi$ ) of Formula I which comprise the corresponding substituents of Formula I described herein. Compounds that are iteratively designed using structural information gleaned from these compounds and which modulate nuclear hormone receptor activity also are of interest. Preferred compounds of Formula I and its derivatives that fit spatially and preferentially into a TR LBD comprise the following substituents:

(i) an R1-substituent comprising an anionic group that interacts with a side chain nitrogen atom of an arginine corresponding to a residue from the group Arg228, Arg262, and Arg266 of human TR- $\alpha$ , and Arg282, Arg316 and Arg320 of human TR- $\beta$ , where the anionic group is about 1.7-4.0Å from the nitrogen atom;

- 1 (ii) an R2-substituent comprising a hydrophobic or hydrophilic group that fits  
2 spacially into the TR LBD;
- 3 (iii) an R3-substituent comprising a hydrophobic or hydrophilic group that  
4 interacts with a side chain atom of a serine, alanine and/or isoleucine corresponding to a  
5 residue from the group Ser260, Ala263 and Ile299 of human TR- $\alpha$ , and Ser314, Ala317 and  
6 Ile352 of human TR- $\beta$ , where the hydrophobic or hydrophilic group is about 1.7-4.0Å from  
7 the side chain atom;
- 8 (iv) an R5-substituent comprising a hydrophobic or hydrophilic group that interacts  
9 with a side chain atom of a phenylalanine and/or isoleucine corresponding to a residue from  
10 the group Phe218, Ile221 and Ile222 of human TR- $\alpha$ , and Phe272, Ile275 and Ile276 of  
11 human TR- $\beta$ , where the hydrophobic or hydrophilic group is about 1.7-4.0Å from the side  
12 chain atom;
- 13 (v) an R6-substituent comprising a hydrophobic or hydrophilic group that fits  
14 spacially into the TR LBD;
- 15 (vi) an X-substituent comprising a hydrophobic or hydrophilic group that interacts  
16 with a side chain atom of a leucine corresponding to a residue from the group Leu276 and  
17 Leu292 of human TR- $\alpha$ , and Leu 330 and Leu346 of human TR- $\beta$ , where the hydrophobic  
18 or hydrophilic group is about 1.7-4.0Å from the side chain atom;
- 19 (vii) an R2'-substituent comprising a hydrophobic or hydrophilic group that fits  
20 spacially into the TR LBD;
- 21 (viii) an R3'-substituent comprising a hydrophobic group that interacts with a side  
22 chain atom of a phenylalanine, glycine and/or methionine corresponding to a residue from  
23 the group Phe215, Gly290, and Met388 of human TR- $\alpha$ , and Phe269, Gly344, Met442 of  
24 human TR- $\beta$ , where the hydrophobic group is about 1.7-4.0Å from the side chain atom;
- 25 (ix) an R4'-substituent comprising an hydrogen bond donor or acceptor group that  
26 interacts with a side chain carbon or nitrogen atom of a histidine corresponding to residue  
27 His381 of human TR- $\alpha$ , and His435 of human TR- $\beta$ , where the hydrogen bond donor or  
28 acceptor group is about 1.7-4.0Å from the side chain atom;
- 29 (x) an R5'-substituent comprising a hydrophobic or hydrophilic group that fits  
30 spacially into the TR LBD;

(xi) and R<sub>6</sub>'-substituent comprising a hydrophobic or hydrophilic group that fits spacially into the TR LBD; and

where the compound is other than thyronine (T3), triiodothyronine (T4) or other thyronine-like compounds previously known and used in a TR treatment method, such as those referenced in Appendix I.

Examples of such substituents include the following:

where R<sub>1</sub> is

-O-CH<sub>2</sub>CO<sub>2</sub>H, -NHCH<sub>2</sub>CO<sub>2</sub>H, -CO<sub>2</sub>H, -CH<sub>2</sub>CO<sub>2</sub>H, -CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H, -CH<sub>2</sub>CH<sub>2</sub>CH<sub>2</sub>CO<sub>2</sub>H, -CH<sub>2</sub>CH(NH<sub>2</sub>)CO<sub>2</sub>H, -CH<sub>2</sub>CH[NHCOCH<sub>2</sub>φ]<sub>2</sub>CO<sub>2</sub>H, -CH<sub>2</sub>CH[NHCO(CH<sub>2</sub>)<sub>15</sub>CH<sub>3</sub>]CO<sub>2</sub>H, -CH<sub>2</sub>CH[NH-FMOC]CO<sub>2</sub>H, -CH<sub>2</sub>CH[NH-tBOC]CO<sub>2</sub>H, or a carboxylate connected to the ring with a 0 to 3 carbon linker,

-PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CH<sub>2</sub>PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CHNH<sub>2</sub>PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CH[NHCOCH<sub>2</sub>φ]<sub>2</sub>PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CH[NHCO(CH<sub>2</sub>)<sub>15</sub>CH<sub>3</sub>]PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CH[NH-FMOC]PO<sub>3</sub>H<sub>2</sub>, -CH<sub>2</sub>CH[NH-tBOC]PO<sub>3</sub>H<sub>2</sub>, or a phosphate or phosphonate connected to the ring with a 0 to 3 carbon linker, -SO<sub>3</sub>H, -CH<sub>2</sub>SO<sub>3</sub>H, -CH<sub>2</sub>CH<sub>2</sub>SO<sub>3</sub>H, -CH<sub>2</sub>CHNH<sub>2</sub>SO<sub>3</sub>H, -CH<sub>2</sub>CH[NHCOCH<sub>2</sub>φ]<sub>2</sub>SO<sub>3</sub>H, -CH<sub>2</sub>CH[NHCO(CH<sub>2</sub>)<sub>15</sub>CH<sub>3</sub>]SO<sub>3</sub>H, -CH<sub>2</sub>CH[NH-FMOC]SO<sub>3</sub>H, -CH<sub>2</sub>CH[NH-tBOC]SO<sub>3</sub>H, or a sulfate or sulfite connected to the ring with a 0 to 3 carbon linker, or acts as the functional equivalent of CH<sub>2</sub>CH(NH<sub>2</sub>)CO<sub>2</sub>H of T3 in the molecular recognition domain when bound to a TR, wherein R<sub>1</sub> can be optionally substituted with an amine,

where R<sub>2</sub> is

H, halogen, CF<sub>3</sub>, OH, NH<sub>2</sub>, SH, CH<sub>3</sub>, -Et, or acts as the functional equivalent of H in the molecular recognition domain when bound to a TR,

where R<sub>3</sub> is

-H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, -N<sub>3</sub>, -SH, -CH<sub>3</sub>, -Et, or acts as the functional equivalent of I in the molecular recognition domain when bound to a TR,

where R<sub>5</sub> is

1 -H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, -N<sub>3</sub>, -SH, -CH<sub>3</sub>, -Et, or acts as the functional  
2 equivalent of I in the molecular recognition domain when bound to a TR, and R<sub>3</sub> can  
3 be identical to R<sub>5</sub>,

4 where R<sub>6</sub> is

5 -H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, -SH, -CH<sub>3</sub>, or acts as the functional equivalent of H  
6 in the molecular recognition domain when bound to a TR, and R<sub>2</sub> can be identical to  
7 R<sub>6</sub>,

8 where R<sub>2</sub>' is

9 -H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, -N<sub>3</sub>, -SH, -CH<sub>3</sub>, -Et, or acts as the functional  
10 equivalent of H in the molecular recognition domain when bound to a TR,  
11 where R<sub>3</sub>' is any hydrophobic group, including

12 halogen, -CF<sub>3</sub>, -SH, alkyl, aryl, 5- or 6-membered heterocyclic, cyano, or acts as the  
13 functional equivalent of I in the molecular recognition domain when bound to a TR,  
14 where R<sub>4</sub>' is

15 -H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, NH<sub>3</sub>, -N(CH<sub>3</sub>)<sub>3</sub>, carboxylate, phosphonate, phosphate  
16 or sulfate, -SH, -CH<sub>3</sub>, -Et, or alkyl, aryl or 5- or 6-membered heterocyclic aromatic  
17 attached through urea or carbamate linkages to O or N or S at the R<sub>4</sub>' position, or  
18 acts as the functional equivalent of OH in the molecular recognition domain when  
19 bound to a TR,

20 where R<sub>5</sub>' is

21 -H, -OH, -NH<sub>2</sub>, -N(CH<sub>3</sub>)<sub>2</sub>, -SH, -NH<sub>3</sub>, -N(CH<sub>3</sub>)<sub>3</sub>, carboxylate, phosphonate, phosphate,  
22 sulfate, branched or straight chain alkyl having 1 to 9 carbons, substituted or  
23 unsubstituted aryl, wherein said substituted aryl is substituted with halogen or 1 to 5  
24 carbon alkyl and wherein said aryl is optionally connected to the ring by a -CH<sub>2</sub>-,  
25 aromatic heterocycle having 5 to 6 atoms, wherein said heterocycle may be substituted  
26 with one or more groups selected from -OH, -NH<sub>2</sub>, -SH, -NH<sub>3</sub>, -N(CH<sub>3</sub>)<sub>3</sub>,  
27 carboxylate, phosphonate, phosphate or sulfate, heteroalkyl, arylalkyl, heteroaryl  
28 alkyl, polyaromatic, or polyheteroaromatic, wherein said R<sub>5</sub>' may be substituted with  
29 polar or charged groups,

30 where R<sub>6</sub>' is

1 -H, halogen, -CF<sub>3</sub>, -OH, -NH<sub>2</sub>, -SH, -CH<sub>3</sub>, -Et, or acts as the functional equivalent of  
2 H in the molecular recognition domain when bound to a TR,  
3 where X is

4 O, S, SO<sub>2</sub>, NH, NR<sub>7</sub>, CH<sub>2</sub>, CHR<sub>7</sub>, CR<sub>7</sub>R<sub>7</sub>, wherein R<sub>7</sub> is alkyl, aryl or 5- or  
5 6-membered heterocyclic aromatic,

6 and where the TR LBD ligand has an apparent K<sub>d</sub> for binding TR LBD of 1 μM or less.

7 Of particular interest are the class of compounds according to Formula I having the  
8 following substituents: where R<sub>1</sub> is carboxylate, phosphonate, phosphate or sulfite and is  
9 connected to the ring with a 0 to 3 carbon linker, R<sub>2</sub> is H, R<sub>3</sub> is -I, -Br, or -CH<sub>3</sub>, R<sub>5</sub> is -I, -  
10 Br, or -CH<sub>3</sub>, R<sub>6</sub> is H, R<sub>2</sub>' is H, R<sub>3</sub>' is -I, -Br, -CH<sub>3</sub>, -iPr, -phenyl, benzyl, or 5- or 6-  
11 membered ring heterocycles, R<sub>4</sub>' is -OH, -NH<sub>2</sub>, and -SH, R<sub>5</sub>' is -H, -OH, -NH<sub>2</sub>, -N(CH<sub>3</sub>)<sub>2</sub> -  
12 SH -NH<sub>3</sub>, -N(CH<sub>3</sub>)<sub>3</sub>, carboxylate, phosphonate, phosphate, sulfate, branched or straight chain  
13 alkyl having 1 to 9 carbons, substituted or unsubstituted aryl, wherein said substituted aryl is  
14 substituted with halogen or 1 to 5 carbon alkyl and wherein said aryl is optionally connected  
15 to the ring by a -CH<sub>2</sub>-, aromatic heterocycle having 5 to 6 atoms, wherein said heterocycle  
16 may be substituted with one or more groups selected from -OH, -NH<sub>2</sub>, -SH, -NH<sub>3</sub>, -N(CH<sub>3</sub>)<sub>3</sub>,  
17 carboxylate, phosphonate, phosphate or sulfate, heteroalkyl, arylalkyl, heteroaryl alkyl,  
18 polyaromatic, or polyheteroaromatic, wherein said R<sub>5</sub>' may be substituted with polar or  
19 charged groups, and R<sub>6</sub>' is H.

20 The present invention also includes a method for identifying a compound capable of  
21 selectively modulating the activity of a TR isoform. By "modulating" is intended increasing  
22 or decreasing activity of a TR. By "TR isoform" is intended TR proteins encoded by  
23 subtype and variant TR genes. This includes TR-α and TR-β isoforms encoded by different  
24 genes (e.g., *thra* and *thrb*) and variants of the same genes (e.g., *thrb1* and *thrb2*). The  
25 method comprises the steps of modeling test compounds that fit spacially and preferentially  
26 into a TR LBD isoform of interest using an atomic structural model of a TR LBD isoform  
27 bound to a test compound, screening the test compounds in a biological assay for TR isoform  
28 activity characterized by binding of a test compound to a TR LBD isoform, and identifying a  
29 test compound that selectively modulates the activity of a TR isoform. By "modeling" is  
30 intended quantitative and qualitative analysis of receptor-ligand structure/function based on  
31 three-dimensional structural information and receptor-ligand interaction models. This

1 includes conventional numeric-based molecular dynamic and energy minimization models,  
2 interactive computer graphic models, modified molecular mechanics models, distance  
3 geometry and other structure-based constraint models. Modeling is preferably performed  
4 using a computer and may be further optimized using known methods.

5 For selectively modulating activity of a TR isoform, such as TR- $\alpha$  or TR- $\beta$ , a  
6 sufficient amount of a compound that fits spatially and preferentially into TR LBD isoform is  
7 provided *in vitro* or *in vivo* to achieve the desired end result. TR- $\alpha$  isoform selectivity can  
8 be accomplished with a compound comprising an anionic group that interacts with an oxygen  
9 or carbon of a serine residue corresponding to Ser277 of human TR- $\alpha$ , where the anionic  
10 group is about 1.7-4.0Å from the side chain atom. TR- $\beta$  isoform selectivity can be  
11 accomplished with a compound comprising an anionic group that interacts with the side chain  
12 nitrogen of an asparagine corresponding to Asn331 of human TR- $\beta$ , where the anionic group  
13 is about 1.7-4.0Å from the side chain nitrogen atom.

14 The present invention further includes a method for identifying a TR agonist or  
15 antagonist ligand by providing the atomic coordinates of a TR LBD to a computerized  
16 modeling system, modeling ligands which fit spacially into the TR LBD, and identifying in a  
17 biological assay for TR activity a ligand which increases or decreases the activity of the TR.  
18

19 The invention also involves a method for increasing receptor selectivity of a  
20 compound of Formula I or derivatives thereof for a TR-type receptor versus other nuclear  
21 receptors by selecting a compound that interacts with conformationally constrained residues  
22 of a TR LBD that are conserved among TR isoforms. "Conformationally constrained" is  
23 intended to refer to the three-dimensional structure of a chemical or moiety thereof having  
24 certain rotations about its bonds fixed by various local geometric and physical-chemical  
25 constraints. In designing and selecting compounds having increased specificity for TRs  
26 compared to other nuclear receptors, the following methods of the invention can be used.  
27 One method involves comparing atomic models of a first TR LBD isoform bound to a  
28 compound with a second TR LBD isoform bound to the same compound, identifying atoms  
29 of the TR LBD and compounds which interact, and designing or selecting a compound that  
30 interacts with TR LBD residues comprising a conformationally constrained structural feature  
31 that is conserved between the TR LBD isoforms. Another method relates to comparing a  
first TR LBD complexed with a first compound to a second TR LBD complexed with a

1 second compound having one or more different substituents compared to the first compound,  
2 identifying atoms of the TR LBD and compounds which interact, and designing or selecting  
3 compounds that interact with TR LBD residues comprising a conformationally constrained  
4 structural feature that is conserved between the TR LBD isoforms. The methods also  
5 facilitate identification of structural and conformationally constrained interactions that are  
6 conserved between compounds that bind to a TR LBD. The methods are exemplified by  
7 comparing atomic models of a first TR LBD isoform complexed with a first compound of  
8 Formula I to a second TR LBD isoform complexed with the first compound, or a second  
9 compound of Formula I having different substituents than the first compound. For example,  
10 a TR- $\alpha$  LBD bound to a natural hormone such as T3 is compared to a TR- $\beta$  LBD bound to  
11 an organic thyronine-like compound such as GC-1. Conserved contacts are identified which  
12 are made between atoms of the different compounds and atoms of the TR LBDs, and the  
13 fiducial and adjustable components identified. Compounds selective for TR are identified in  
14 a biological assay for TR activity that assays for selective binding to a TR and/or TR LBD -  
15 compared to other nuclear receptors. Conventional assays for TR and other nuclear  
16 receptors may be conducted in parallel or serially, including those assays described herein.  
17 Automatable methods are preferred. The methods facilitate design and selection of  
18 compounds comprising cyclic carbon and substituent atoms that interact with a constrained  
19 side chain and/or main chain atom of a TR LBD residue.

20 In another aspect of the invention, the methods described herein are useful for  
21 selecting peptides, peptidomimetics or synthetic molecules that modulate TR activity.  
22 Methods of the invention also find use in characterizing structure/function relationships of  
23 natural and synthetic TR-ligands. Molecules of particular interest are new thyronine-like  
24 compounds other than T3, T4 and other thyronine-like compounds previously known and  
25 used for treating TR-related disorders. New compounds of the invention include those which  
26 bind to a TR LBD isoform with greater affinity than T3 or T4 and those which exhibit  
27 isoform-specific binding affinity.

## 28 29 **APPLICABILITY TO NUCLEAR RECEPTORS**

30 The present invention, particularly the computational methods, can be used to design  
31 drugs for a variety of nuclear receptors, such as receptors for glucocorticoids (GRs),



1 androgens (ARs), mineralocorticoids (MRs), progestins (PRs), estrogens (ERs), thyroid  
2 hormones (TRs), vitamin D (VDRs), retinoid (RARs and RXRs), icosanoid (IRs), and  
3 peroxisomes (XPARS and peroxisomal proliferators (PPAP)). The present invention can also  
4 be applied to the "orphan receptors," as they are structurally homologous in terms of  
5 modular domains and primary structure to classic nuclear receptors, such as steroid and  
6 thyroid receptors. The amino acid homologies of orphan receptors with other nuclear  
7 receptors ranges from very low ( $<15\%$ ) to in the range of 35% when compared to rat  
8 RAR $\alpha$  and human TR- $\beta$  receptors, for example. In addition, as is revealed by the X-ray  
9 crystallographic structure of the TR and structural analysis disclosed herein, the overall  
10 folding of liganded superfamily members is likely to be similar. Although ligands have not  
11 been identified with orphan receptors, once such ligands are identified one skilled in the art  
12 will be able to apply the present invention to the design and use of such ligands, as their  
13 overall structural modular motif will be similar to other nuclear receptors described herein.

#### 14 15 *Modular Functional Domains Of Nuclear receptors*

16 The present invention will usually be applicable to all nuclear receptors, as discussed  
17 herein, in part, to the patterns of nuclear receptor activation, structure and modulation that  
18 have emerged as a consequence of determining the three dimensional structures of nuclear  
19 receptors with different ligands bound, notably the three dimensional structures or  
20 crystallized protein structure of the ligand binding domains for TR- $\alpha$  and TR- $\beta$ . Proteins of  
21 the nuclear receptor superfamily display substantial regions of amino acid homology, as  
22 described herein and known in the art see FIG. 2. Members of this family display an  
23 overall structural motif of three modular domains (which is similar to the TR three modular  
24 domain motif):

- 25 1) a variable amino-terminal domain;
- 26 2) a highly conserved DNA-binding domain (DBD); and
- 27 3) a less conserved carboxyl-terminal LBD.

28 The modularity of this superfamily permits different domains of each protein to separately  
29 accomplish different functions, although the domains can influence each other. The separate  
30 function of a domain is usually preserved when a particular domain is isolated from the  
31 remainder of the protein. Using conventional protein chemistry techniques a modular domain

1 can sometimes be separated from the parent protein. Using conventional molecular biology  
2 techniques each domain can usually be separately expressed with its original function intact  
3 or chimerics of two different nuclear receptors can be constructed, wherein the chimerics  
4 retain the properties of the individual functional domains of the respective nuclear receptors  
5 from which the chimerics were generated.

6 **FIG. 2** provides a schematic representation of family member structures, indicating  
7 regions of homology within family members and functions of the various domains.  
8

### 9 *Amino Terminal Domain*

10 The amino terminal domain is the least conserved of the three domains and varies  
11 markedly in size among nuclear receptor superfamily members. For example, this domain  
12 contains 24 amino acids in the VDR and 603 amino acids in the MR. This domain is  
13 involved in transcriptional activation and in some cases its uniqueness may dictate selective  
14 receptor-DNA binding and activation of target genes by specific receptor isoforms. This  
15 domain can display synergistic and antagonistic interactions with the domains of the LBD.  
16 For example, studies with mutated and/or deleted receptors show positive cooperativity of the  
17 amino and carboxy terminal domains. In some cases, deletion of either of these domains will  
18 abolish the receptor's transcriptional activation functions.  
19

### 20 *DNA-Binding Domain*

21 The DBD is the most conserved structure in the nuclear receptor superfamily. It  
22 usually contains about 70 amino acids that fold into two zinc finger motifs, wherein a zinc  
23 ion coordinates four cysteines. DBDs contain two perpendicularly oriented  $\alpha$ -helixes that  
24 extend from the base of the first and second zinc fingers. The two zinc fingers function in  
25 concert along with non-zinc finger residues to direct nuclear receptors to specific target sites  
26 on DNA and to align receptor homodimer or heterodimer interfaces. Various amino acids in  
27 DBD influence spacing between two half-sites (usually comprised of six nucleotides) for  
28 receptor dimer binding. For example, GR subfamily and ER homodimers bind to half-sites  
29 spaced by three nucleotides and oriented as palindromes. The optimal spacings facilitate  
30 cooperative interactions between DBDs, and D box residues are part of the dimerization

1 interface. Other regions of the DBD facilitate DNA-protein and protein-protein interactions  
2 required for RXR homodimerization and heterodimerization on direct repeat elements.

3 The LBD may influence the DNA binding of the DBD, and the influence can also be  
4 regulated by ligand binding. For example, TR ligand binding influences the degree to which  
5 a TR binds to DNA as a monomer or dimer. Such dimerization also depends on the spacing  
6 and orientation of the DNA half sites. The receptors also can interact with other proteins  
7 and function to regulate gene expression.

8 The nuclear receptor superfamily has been subdivided into two subfamilies: 1) GR  
9 (GR, AR, MR and PR) and 2) TR (TR, VDR, RAR, RXR, and most orphan receptors) on  
10 the basis of DBD structures, interactions with heat shock proteins (hsp), and ability to form  
11 heterodimers. GR subgroup members are tightly bound by hsp in the absence of ligand,  
12 dimerize following ligand binding and dissociation of hsp, and show homology in the DNA  
13 half sites to which they bind. These half sites also tend to be arranged as palindromes. TR  
14 subgroup members tend to be bound to DNA or other chromatin molecules when unliganded,  
15 can bind to DNA as monomers and dimers, but tend to form heterodimers, and bind DNA  
16 elements with a variety of orientations and spacings of the half sites, and also show  
17 homology with respect to the nucleotide sequences of the half sites. By this classification,  
18 ER does not belong to either subfamily, since it resembles the GR subfamily in hsp  
19 interactions, and the TR subfamily in nuclear localization and DNA-binding properties.  
20

### 21 *Ligand Binding Domain*

22 The LBD is the second most highly conserved domain in these receptors. Whereas  
23 integrity of several different LBD sub-domains is important for ligand binding, truncated  
24 molecules containing only the LBD retain normal ligand-binding activity. This domain also  
25 participates in other functions, including dimerization, nuclear translocation and  
26 transcriptional activation, as described herein. Importantly, this domain binds the ligand and  
27 undergoes ligand-induced conformational changes as detailed herein.

28 Most members of the superfamily, including orphan receptors, possess at least two  
29 transcription activation subdomains, one of which is constitutive and resides in the amino  
30 terminal domain (AF-1), and the other of which (AF-2 (also referenced as TAU 4)) resides  
31 in the ligand-binding domain whose activity is regulated by binding of an agonist ligand.

1 The function of AF-2 requires an activation domain (also called transactivation domain) that  
2 is highly conserved among the receptor superfamily (approximately amino acids 1005 to  
3 1022). Most LBDs contain an activation domain. Some mutations in this domain abolish  
4 AF-2 function, but leave ligand binding and other functions unaffected. Ligand binding  
5 allows the activation domain to serve as an interaction site for essential co-activator proteins  
6 that function to stimulate (or in some cases, inhibit) transcription.

7 For example, Shibata, H., *et al.* (*Recent Progress in Hormone Res.* 52:141-164  
8 (1997)) has reviewed the role of co-activators and co-repressors in steroid/thyroid hormone  
9 receptor systems. Steroid receptor co-activator-one (SRC-1) appears to be a general co-  
10 activator for all AF-2 domain containing receptors tested. SRC-1 enhances transactivation of  
11 steroid hormone-dependent target genes. Other putative co-activators have been reported,  
12 including the SRC-1 related proteins, TIF-2 and GRIP-1, and other putative unrelated co-  
13 activators such as ARA-70, Trip 1, RIP-140, and TIF-1. In addition another co-activator  
14 CREB-binding protein (CBP) has been shown to enhance receptor-dependent target gene  
15 transcription. CBP and SRC-1 interact and synergistically enhance transcriptional activation  
16 by the ER and PR. A ternary complex of CBP, SRC-1, and liganded receptors may form to  
17 increase the rate of hormone-responsive gene transcription. Co-repressors, such as SMRT  
18 and N-CoR, for TR and RAR, have been identified that also contribute to the silencing  
19 function of unliganded TR. The unliganded TR and RAR have been shown to inhibit basal  
20 promoter activity; this silencing of target gene transcription by unliganded receptors is  
21 mediated by these co-repressors. The collective data suggests that upon binding of agonist,  
22 the receptor changes its conformation in the ligand-binding domain that enables recruitment  
23 of co-activators, which allows the receptor to interact with the basal transcriptional  
24 machinery more efficiently and to activate transcription. In contrast, binding of antagonists  
25 induces a different conformational change in the receptor. Although some antagonist-bound  
26 receptors can dimerize and bind to their cognate DNA elements, they fail to dislodge the  
27 associated co-repressors, which results in a nonproductive interaction with the basal  
28 transcriptional machinery. Similarly, the TR and RAR associate with co-repressors in the  
29 absence of ligand, thereby resulting in a negative interaction with the transcriptional  
30 machinery that silences target gene expression. In the case of mixed agonist/antagonists,  
31 such as 4-hydroxytamoxifen, activation of gene transcription may depend on the relative ratio

1 of co-activators and co-repressors in the cell or cell-specific factors that determine the  
2 relative agonistic or antagonistic potential of different compounds. These co-activators and  
3 co-repressors appear to act as an accelerator and/or a brake that modulates transcriptional  
4 regulation of hormone-responsive target gene expression.

5 The carboxy-terminal activation subdomain, as described herein is in close three  
6 dimensional proximity in the LBD to the ligand, so as to allow for ligands bound to the LBD  
7 to coordinate (or interact) with amino acid(s) in the activation subdomain. As described  
8 herein, the LBD of a nuclear receptor can be expressed, crystallized, its three dimensional  
9 structure determined with a ligand bound (either using crystal data from the same receptor or  
10 a different receptor or a combination thereof), and computational methods used to design  
11 ligands to its LBD, including ligands that contain an extension moiety that coordinates the  
12 activation domain of the nuclear receptor.

13 Once a computationally designed ligand (CDL) is synthesized as described herein and  
14 known in the art, it can be tested using assays to establish its activity as an agonist, partial  
15 agonist or antagonist, and affinity, as described herein. After such testing, the CDLs can be  
16 further refined by generating LBD crystals with a CDL bound to the LBD. The structure of  
17 the CDL can then be further refined using the chemical modification methods described  
18 herein for three dimensional models to improve the activity or affinity of the CDL and make  
19 second generation CDLs with improved properties, such as that of a super agonist or  
20 antagonist described herein. Agonist and antagonist ligands also can be selected that  
21 modulate nuclear receptor responsive gene transcription through altering the interaction of  
22 co-activators and co-repressors with their cognate nuclear hormone receptor. For example,  
23 CDL agonists can be selected that block or dissociate the co-repressor from interaction with  
24 the receptor, and/or which promote binding or association of the co-activator. CDL  
25 antagonists can be selected that block co-activator interaction and/or promote co-repressor  
26 interaction with the target receptor. Selection can be done in binding assays that screen for  
27 CDLs having the desired agonist or antagonist properties. Suitable assays for such screening  
28 are described herein and in Shibata, H., *et al.* (*Recent Prog. Horm. Res.* 52:141-164  
29 (1997)); Tagami, T., *et al.* (*Mol. Cell Biol.* 17(5):2642-2648 (1997)); Zhu, XG., *et al.* (*J.*  
30 *Biol. Chem.* 272(14):9048-9054 (1997)); Lin, B.C., *et al.* (*Mol. Cell Biol.* 17(10):6131-6138  
31 (1997)); Kakizawa, T., *et al.* (*J. Biol. Chem.* 272(38):23799-23804 (1997)); and Chang, K.

1 H., et al. (*Proc. Natl. Acad. Sci. USA* 94(17):9040-9045 (1997)), which references are  
2 incorporated herein in their entirety by reference.

#### 4 NUCLEAR RECEPTOR ISOFORMS

5 The present invention also is applicable to generating new synthetic ligands to  
6 distinguish nuclear receptor isoforms. As described herein, CDLs can be generated that  
7 distinguish between binding isoforms, thereby allowing the generation of either tissue specific  
8 or function specific synthetic ligands. For instance, GR subfamily members have usually one  
9 receptor encoded by a single gene, although are exceptions. For example, there are two PR  
10 isoforms, A and B, translated from the same mRNA by alternate initiation from different  
11 AUG codons. There are two GR forms, one of which does not bind ligand. This method is  
12 especially applicable to the TR subfamily which usually has several receptors that are  
13 encoded by at least two (TR:  $\alpha$ ,  $\beta$ ) or three (RAR, RXR, and PPAR:  $\alpha$ ,  $\beta$ ,  $\gamma$ ) genes or have  
14 alternate RNA splicing and such an example for TR is described herein.

#### 16 NUCLEAR RECEPTOR CRYSTALS

17 The invention provides for crystals made from nuclear receptor ligand binding  
18 domains with the ligand bound to the receptor. As exemplified in the Examples, TRs are  
19 crystallized with a ligand bound to it. Crystals are made from purified nuclear receptor  
20 LBDs that are usually expressed by a cell culture, such as *E. coli*. Preferably, different  
21 crystals (co-crystals) for the same nuclear receptor are separately made using different  
22 ligands, such as a naturally occurring ligand and at least one bromo- or iodo- substituted  
23 synthetic ligand that acts as an analog or antagonist of the naturally occurring ligand. Such  
24 bromo- and iodo- substitutions act as heavy atom substitutions in nuclear receptor ligands and  
25 crystals of nuclear receptor proteins. This method has the advantage for phasing of the  
26 crystal in that it bypasses the need for obtaining traditional heavy metal derivatives. After  
27 the three dimensional structure is determined for the nuclear receptor LBD with its ligand  
28 bound, the three dimensional structure can be used in computational methods to design a  
29 synthetic ligand for the nuclear receptor and further activity structure relationships can be  
30 determined through routine testing using the assays described herein and known in the art.

## 1 *Expression and Purification of other Nuclear Receptor LBD Structures*

2 High level expression of nuclear receptor LBDs can be obtained by the techniques  
3 described herein as well as others described in the literature. High level expression in E.  
4 coli of ligand binding domains of TR and other nuclear receptors, including members of the  
5 steroid/thyroid receptor superfamily, such as the receptors ER, AR, MR, PR, RAR, RXR  
6 and VDR can also be achieved. Yeast and other eukaryotic expression systems can be used  
7 with nuclear receptors that bind heat shock proteins as these nuclear receptors are generally  
8 more difficult to express in bacteria, with the exception of ER, which can be expressed in  
9 bacteria. Representative nuclear receptors or their ligand binding domains have been cloned  
10 and sequenced: human RAR- $\alpha$ , human RAR- $\gamma$ , human RXR- $\alpha$ , human RXR- $\beta$ , human  
11 PPAR- $\alpha$ , human PPAR- $\beta$ , human PPAR- $\gamma$ , human VDR, human ER (as described in  
12 Seielstad *et al.*, *Molecular Endocrinology*, vol 9:647-658 (1995), incorporated herein by  
13 reference), human GR, human PR, human MR, and human AR. The ligand binding domain  
14 of each of these nuclear receptors has been identified and is shown in ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup>. Using the  
15 information in ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup> in conjunction with the methods described herein and known in the  
16 art, one of ordinary skill in the art could express and purify LBDs of any of the nuclear  
17 receptors, including those illustrated in ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup>, bind it to an appropriate ligand, and  
18 crystallize the nuclear receptor's LBD with a bound ligand.

19 ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup> is an alignment of several members of the steroid/thyroid hormone receptor  
20 superfamily that indicates the amino acids to be included in a suitable expression vector.

21 Extracts of expressing cells are a suitable source of receptor for purification and  
22 preparation of crystals of the chosen receptor. To obtain such expression, a vector is  
23 constructed in a manner similar to that employed for expression of the rat TR alpha (Apriletti  
24 *et al.* *Protein Expression and Purification*, 6:363-370 (1995), herein incorporated by  
25 reference). The nucleotides encoding the amino acids encompassing the ligand binding  
26 domain of the receptor to be expressed, for example the estrogen receptor ligand binding  
27 domain (hER-LBD) (corresponding to R at position 725 to L at position 1025 as standardly  
28 aligned as shown in the ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup>), are inserted into an expression vector such as the one  
29 employed by Apriletti *et al.* (1995). For the purposes of obtaining material that will yield  
30 good crystals it is preferable to include at least the amino acids corresponding to human TR-  
31  $\beta$  positions 725 to 1025. Stretches of adjacent amino acid sequences may be included if

1 more structural information is desired. Thus, an expression vector for the human estrogen  
2 receptor can be made by inserting nucleotides encoding amino acids from position 700 to the  
3 c-terminus at position 1071. Such a vector gives high yield of receptor in E. coli that can  
4 bind hormone (Seielstad *et al. Molecular Endocrinology* 9:647-658 (1995)). However, the c-  
5 terminal region beyond position 1025 is subject to variable proteolysis and can  
6 advantageously be excluded from the construct, this technique of avoiding variable  
7 proteolysis can also be applied to other nuclear receptors.

### 9 *TR- $\alpha$ And TR- $\beta$ As Examples of Nuclear receptor LBD Structure and Function*

#### 10 *TR Expression, Purification And Crystallization*

11 As an example of nuclear receptor structure of the ligand binding domain the  $\alpha$ - and  
12  $\beta$ - isoforms of TR are crystallized from proteins expressed from expression constructs,  
13 preferably constructs that can be expressed in E. coli. Other expression systems, such as  
14 yeast or other eukaryotic expression systems can be used. For the TR, the LBD can be  
15 expressed without any portion of the DBD or amino-terminal domain. Portions of the DBD  
16 or amino-terminus can be included if further structural information with amino acids adjacent  
17 the LBD is desired. Generally, for the TR the LBD used for crystals will be less than 300  
18 amino acids in length. Preferably, the TR LBD will be at least 150 amino acids in length,  
19 more preferably at least 200 amino acids in length, and most preferably at least 250 amino  
20 acids in length. For example the LBD used for crystallization can comprise amino acids  
21 spanning from Met 122 to Val 410 of the rat TR- $\alpha$ , Glu 202 to Asp 461 of the human TR- $\beta$ .

22 Typically TR LBDs are purified to homogeneity for crystallization. Purity of TR  
23 LBDs is measured with sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-  
24 PAGE), mass spectrometry (MS) and hydrophobic high performance liquid chromatography  
25 (HPLC). The purified TR for crystallization should be at least 97.5 % pure or 97.5%,  
26 preferably at least 99.0% pure or 99.0% pure, more preferably at least 99.5% pure or  
27 99.5% pure.

28 Initially purification of the unliganded receptor can be obtained by conventional  
29 techniques, such as hydrophobic interaction chromatography (HPLC), ion exchange  
30 chromatography (HPLC), and heparin affinity chromatography.



1 To achieve higher purification for improved crystals of nuclear receptors, especially  
2 the TR subfamily and TR, it will be desirable to ligand shift purify the nuclear receptor using  
3 a column that separates the receptor according to charge, such as an ion exchange or  
4 hydrophobic interaction column, and then bind the eluted receptor with a ligand, especially  
5 an agonist. The ligand induces a change in the receptor's surface charge such that when re-  
6 chromatographed on the same column, the receptor then elutes at the position of the liganded  
7 receptor are removed by the original column run with the unliganded receptor. Usually  
8 saturating concentrations of ligand are used in the column and the protein can be  
9 preincubated with the ligand prior to passing it over the column. The structural studies  
10 detailed herein indicate the general applicability of this technique for obtaining super-pure  
11 nuclear receptor LBDs for crystallization.

12 More recently developed methods involve engineering a "tag" such as with histidine  
13 placed on the end of the protein, such as on the amino terminus, and then using a nickle  
14 chelation column for purification, Janknecht R., *Proc. Natl. Acad. Sci. USA*, 88:8972-8976-  
15 (1991) incorporated by reference.

16 To determine the three dimensional structure of a TR LBD, or a LBD from another  
17 member of the nuclear receptor superfamily, it is desirable to co-crystallize the LBD with a  
18 corresponding LBD ligand. In the case of TR LBD, it is preferable to separately co-  
19 crystalize it with ligands such as T3, IpBr and Dimit that differ in the heavy atoms which  
20 they contain. Other TR ligands such as those encompassed by Formula 1 described herein  
21 and known in the prior art, can also be used for the generation of co-crystals of TR LBD and  
22 TR ligands. Of the compounds encompassed by Formula 1 it is generally desirable to use at  
23 least one ligand that has at least one bromo- or iodo- substitution at the  $R_3$ ,  $R_5$ ,  $R_3'$  or  $R_5'$   
24 position, preferably such compounds will be have at least two such substitutions and more  
25 preferably at least 3 such substitutions. As described herein, such substitutions are  
26 advantageously used as heavy atoms to help solve the phase problem for the three  
27 dimensional structure of the TR LBD and can be used as a generalized method of phasing  
28 using a halogen (e.g. I or Br) substituted ligand, especially for nuclear receptors.

29 Typically purified LBD, such as TR LBD, is equilibrated at a saturating concentration  
30 of ligand at a temperature that preserves the integrity of the protein. Ligand equilibration

1 can be established between 2 and 37°C, although the receptor tends to be more stable in the  
2 2-20°C range.

3 Preferably crystals are made with the hanging drop methods detailed herein.  
4 Regulated temperature control is desirable to improve crystal stability and quality.  
5 Temperatures between 4 and 25°C are generally used and it is often preferable to test  
6 crystallization over a range of temperatures. In the case of TR it is preferable to use  
7 crystallization temperatures from 18 to 25°C, more preferably 20 to 23°C, and most  
8 preferably 22°C.

9 Complexes of the TR- $\alpha$  LBD with a variety of agonists, including T<sub>3</sub>, IpBr<sub>2</sub>, Dimit,  
10 and Triac, are prepared with by methods described herein. For example, cocrystals of the  
11 rTR- $\alpha$  LBD, with ligand prebound, are prepared by vapor diffusion at ambient temperature  
12 from 15% 2-methyl-2,4-pentanediol (MPD). The crystals are radiation sensitive, and require  
13 freezing to measure complete diffraction data. On a rotating anode X-ray source, the  
14 crystals diffract to  $\sim 3\text{\AA}$ ; synchrotron radiation extends the resolution limit significantly, to  
15 as high as 2.0 $\text{\AA}$  for T<sub>3</sub> cocrystals. The composition of the thyroid hormone, combined with  
16 the ability to prepare and cocrystallize the receptor complexed with a variety of analogs,  
17 permitted the unusual phasing strategy. This phasing strategy can be applied to the ligands  
18 of the nuclear receptors described therein by generating I and Br substitutions of such  
19 ligands. In this strategy, cocrystals of the TR LBD containing four hormone analogs that  
20 differ at the 3,5, and 3' positions (T<sub>3</sub>, IpBr<sub>2</sub>, Dimit, and Triac) provided isomorphous  
21 derivatives. For this set of analogs, the halogen substituents (2Br and 3I atoms) function as  
22 heavy atoms, while the Dimit cocrystal (3 alkyl groups) acts as the parent. The initial 2.5 $\text{\AA}$   
23 multiple isomorphous replacement/anomalous scattering/density modified electron density  
24 map allowed the LBD to be traced from skeletons created in the molecular graphics program  
25 O5 (Jones, T.A. *et al.*, *ACTA Cryst*, 47:110-119 (1991), incorporated by reference herein).  
26 A model of the LBD was built in four fragments, Arg157-Gly184, Trp186-Gly197, Ser199-  
27 Pro205, and Val210-Phe405, and refined in XPLOR using positional refinement and  
28 simulated annealing protocols. Missing residues were built with the aid of difference  
29 density. The final model was refined to  $R_{\text{cryst}} = 21.8\%$  and  $R_{\text{free}} = 24.4\%$  for data from  
30 15.0 to 2.2 $\text{\AA}$ , see Table 6. The human TR- $\beta$  LBD model was resolved by molecular  
31 replacement of the TR- $\alpha$  LBD coordinates. The structure is based on E202 to D461 with a

1 his-tag at the N-terminus. The final model was refined to  $R_{\text{cryst}} = 25.3\%$  and  $R_{\text{free}} = 28.9\%$   
2 for data from 30.0 to 2.4 Å, see Table 7.

3 This phasing strategy can be applied to the ligands of the nuclear receptors described  
4 herein by generating I and Br substitutions of such ligands.

## 6 **THREE DIMENSIONAL STRUCTURE OF TR LBD**

### 7 *Architecture of TR LBD*

8 As an example of the three dimensional structure of a nuclear receptor, the folding of  
9 the TR- $\alpha_1$  LBD is shown in FIG. 4. The TR- $\alpha$  LBD consists of a single structural domain  
10 packed in three layers, composed of twelve  $\alpha$ -helices, H1-12, and four short  $\beta$ -strands, S1-4,  
11 forming a mixed  $\beta$ -sheet. The buried hormone and three antiparallel  $\alpha$ -helices, H5-6, H9,  
12 and H10, form the central layer of the domain, as shown in FIG. 4. H1, H2, H3 and S1  
13 form one face of the LBD, with the opposite face formed by H7, H8, H11, and H12. The  
14 first 35 amino acids of the N-terminus (Met122-Gln156) are not visible in the electron  
15 density maps. The three dimensional structure of the heterodimeric RXR:TR DNA-binding  
16 domains bound to DNA, amino acids Met 122 - Gln151 of the TR DBD make extensive  
17 contacts with the minor groove of the DNA8. The five disordered amino acids (Arg152-  
18 Gln156), which reside between the last visible residue of the TR DBD and the first visible  
19 residue of the LBD likely represent the effective "hinge" linking the LBD and the DBD in  
20 the intact receptor.

21 The predominantly helical composition and the layered arrangement of secondary  
22 structure is identical to that of the unliganded hRXR $\alpha$ , confirming the existence of a common  
23 nuclear receptor fold between two nuclear receptors.

24 The TR LBD is visible beginning at Arg157, and continues in an extended coil  
25 conformation to the start of H1. A turn of  $\alpha$ -helix, H2, covers the hormone binding cavity,  
26 immediately followed by short  $\beta$ -strand, S1, which forms the edge of the mixed  $\beta$ -sheet,  
27 parallel to S4, the outermost of the three antiparallel strands. The chain is mostly irregular  
28 until H3 begins, antiparallel to H1. H3 bends at Ile221 and Ile222, residues which contact  
29 the ligand. The chain turns almost 90° at the end of H3 to form an incomplete  $\alpha$ -helix, H4.  
30 The first buried core helix, H5-6, follows, its axis altered by a kink near the ligand at Gly  
31 253. The helix is composed of mostly hydrophobic sidechains interrupted by two striking

1 exceptions: Arg262 is solvent inaccessible and interacts with the ligand carboxylate (1-  
 2 substituent), and Glu256 meets Arg329 from H9 and Arg375 from H11 in a polar  
 3 invagination. H5-6 terminates in a short  $\beta$ -strand, S2, of the four strand mixed sheet. S3  
 4 and S4 are joined through a left-handed turn, and further linked by a salt bridge between  
 5 Lys284 and Asp272. Following S4, H7 and H8 form an L, stabilized by a salt bridge  
 6 between Lys268 and Asp277. The turn between H7 and H8 adopts an unusual conformation,  
 7 a result of interaction with ligand and its glycine rich sequence. H9 is the second core helix.  
 8 antiparallel to the neighboring H5-6. Again, two buried polar sidechains are found, Glu315  
 9 and Gln320. Glu315 forms a buried salt bridge with His358 and Arg356. The oxygen of  
 10 Gln320 forms a hydrogen bond with the buried sidechain of His 175. The chain then  
 11 switches back again to form H10, also antiparallel to H9. H11 extends diagonally across the  
 12 full length of the molecule. Immediately after H11, the chain forms a type II turn, at  
 13 approximately  $90^\circ$  to H11. The chain then turns again to form H 12, which packs loosely  
 14 against H3 and H11 as part of the hormone or ligand binding cavity. The final five amino -  
 15 acids at the C-terminus, Glu406 -Val410, are disordered. The architecture of the TR- $\beta$  LBD  
 16 is identical to that of the TR- $\alpha$  LBD, with two significant differences. An additional helix is  
 17 present at the N-terminus (residues Glu202-Ile208), which is part of the DBD, and packs  
 18 antiparallel to H10. Following the helix is a two residue turn (Gly209-His210) continuing  
 19 into an extended coil to the start of H1, as seen in the TR- $\alpha$  LBD. A further difference  
 20 occurs in the irregular conformation adopted between H2 and H3. In the TR- $\alpha$  LBD,  
 21 residue Gly197-Asp211 form a loop that packs against the receptor, contacting helices H7,  
 22 H8, H11, and the loop between H11 and H12. In the TR- $\beta$  LBD, only the ends of the loop  
 23 are ordered, with the stretch Ala253-Lys263 disordered. In addition to these residues, the  
 24 residues of the His-tag at the N-terminus, and the final residue at the C-terminus, Asp461,  
 25 are disordered.

## 27 *TR LBD's Ligand Binding Cavity As An Example Of A Nuclear Receptor's Buried Ligand* 28 *Cavity*

29 The three dimensional structure of the TR LBD leads to the startling finding that  
 30 ligand binding cavity of the LBD is solvent inaccessible when a T3 or its isostere is bound to  
 31 the LBD. This surprising result leads to a new model of nuclear receptor three dimensional

1 structure and function, as further described herein, particularly in the sections elucidating the  
2 computational methods of ligand design and the application of such methods to designing  
3 nuclear receptor synthetic ligands that contain extended positions that prevent normal  
4 activation of the activation domain.

5 Dimit, the ligand bound to the receptor, is an isostere of  $T_3$  and a thyroid hormone  
6 agonist. Therefore the binding of Dimit should reflect that of  $T_3$ , and the Dimit-bound  
7 receptor is expected to be the active conformation of TR. The ligand is buried within the  
8 receptor, providing the hydrophobic core for a subdomain of the protein, as shown in **FIG. 5**  
9 **a and b**. H5-6 and H9 comprise the hydrophobic core for the rest of the receptor.

10 An extensive binding cavity is constructed from several structural elements. The  
11 cavity is enclosed from above by H5-6 (Met 256- Arg266), from below by H7 and H8 and  
12 the intervening loop (Leu287- Ile299), and along the sides by H2 (185-187), by the turn  
13 between S3 and S4 (Leu276-Ser277), by H3 (Phe215-Arg228), by H11 (His381-Met388) and  
14 by H12 (Phe401-Phe405). The volume of the cavity defined by these elements, calculated by  
15 GRASP (Columbia University, USA) (600 Å<sup>3</sup>), is essentially the volume of the hormone  
16 (530 Å<sup>3</sup>). The change in volume can be exploited for ligand design as described herein. The  
17 remaining volume is occupied by water molecules surrounding the amino-propionic acid  
18 substituent. **FIG. 6** depicts various contacts (or interactions) between TR's LBD and the  
19 ligand.

20 The planes of the inner and outer (prime ring) rings of the ligand are rotated from  
21 planarity about 60° with respect to each other, adopting the 3'-distal conformation (in which  
22 the 3' substituent of the outer ring projects down and away from the inner ring). The amino-  
23 propionic acid and the outer phenolic ring assume the transoid conformation, each on  
24 opposite sides of the inner ring. The torsion angle  $\chi_1$  for the amino- propionic acid is 300°.

25 The amino-propionic acid substituent is packed loosely in a polar pocket formed by  
26 side chains from H2, H4 and S3. The carboxylate group forms direct hydrogen bonds with  
27 the guanidium group of Arg228 and the amino N of Ser277. In addition, Arg262, Arg266  
28 and Asn179 interact with the carboxylate through water-mediated hydrogen bonds. The three  
29 arginine residues create a significantly positive local electrostatic potential, which may  
30 stabilize the negative charge of the carboxylate. No hydrogen bond is formed by the amino  
31 nitrogen. The interactions of the amino-propionic acid substituent are consistent with the fact

1 that Triac, which lacks the amino nitrogen, has a binding affinity equal to that of  $T_3$ ,  
2 indicating that the amino nitrogen and longer aliphatic chain of  $T_3$  do not contribute greatly to  
3 binding affinity.

4 The biphenyl ether, in contrast, is found buried within the hydrophobic core. The  
5 inner ring packs in a hydrophobic pocket formed by H3, H5-6, and S3. Pockets for the 3-  
6 and 5-methyl substituents are not completely filled, as expected since the van der Waals  
7 radius of methyl substituent for Dimit is smaller than the iodine substituent provided by the  
8 thyroid hormone  $T_3$ . Such pockets are typically 25 to 100 cubic angstroms (although smaller  
9 pocket for substitutes are contemplated in the 40 to 80 cubic angstrom range) and could be  
10 filled more tightly with better fitting chemical substitutions, as described herein.

11 The outer ring packed tightly in a pocket formed by H3, H5-6, H7, H8, H11 and  
12 H12, and the loop between H7 and H8. The ether oxygen is found in a hydrophobic  
13 environment defined by Phe218, Leu287, Leu276, and Leu292. The absence of a hydrogen  
14 bond to the ether oxygen is consistent with its role in establishing the correct stereochemistry  
15 of the phenyl rings, as suggested by potent binding of hormone analogs with structurally  
16 similar linkages possessing reduced or negligible hydrogen bonding capability. The 3'-  
17 isopropyl substituent contacts Gly290 and 291. The presence of glycine at this position in  
18 the pocket can explain the observed relationship between activity and the size of 3'-  
19 substituents. Activity is highest for 3'-isopropyl, and decreases with added bulk. The only  
20 hydrogen bond in the hydrophobic cavity is formed between the phenolic hydroxyl and  
21 His381 N $\epsilon$ 2. The conformation of His381 is stabilized by packing contacts provided by  
22 Phe405, and Met256.

23 The presence of a 5' substituent larger than hydrogen affects the binding affinity for  
24 hormone. The more abundant thyroid hormone, 3,5,3',5'-tetraiodo-L-thyronine ( $T_4$ ), contains  
25 an iodine at this position, and binds the receptor with 2% of the affinity of  $T_3$ . The structure  
26 suggests that discrimination against  $T_4$  is accomplished through the combination of steric  
27 conflict by Met256 and possibly the constraints imposed by the geometry of the hydrogen  
28 bond from His381 to the phenolic hydroxyl. The 5' position is a preferred location for  
29 introducing a chemical modification of C-H at the 5' of  $T_3$  or and TR agonist, as described  
30 herein, that produces an extension from the prime ring and results in the creation of an  
31 antagonist or partial agonist.

1 Deletion and antibody competition studies suggest the involvement of residues Pro162  
2 to Val202 in ligand binding. The region does not directly contact hormone in the bound  
3 structure, although H2 packs against residues forming the polar pocket that interacts with the  
4 amino-propionic acid group. One role for H2, then, is to stabilize these residues in the bound  
5 state, H2, with  $\beta$ -strands S3 and S4, might also represent a prevalent entry point for ligand,  
6 since the amino-propionic acid of the ligand is oriented toward this region. Studies of  
7 receptor binding to  $T_3$  affinity matrices demonstrate that only a linkage to the amino-  
8 propionic acid is tolerated, suggesting that steric hindrance present in other linkages prevent  
9 binding. Furthermore, the crystallographic temperature factors suggest the coil and  $\beta$ -strand  
10 region is most flexible part of the domain FIG. 7. Participation of this region, part of the  
11 hinge domain between the DBD and LBD, in binding hormone may provide structural means  
12 for ligand binding to influence DNA binding, since parts of the Hinge domain contact DNA.

#### 13 14 *TR LBD Transcriptional Activation Helix As An Example Of A Nuclear Receptor* 15 *Activation Domain*

16 In addition to the startling finding that the ligand binding cavity is solvent inaccessible  
17 when loaded with a ligand, the activation helix of TR LBD presents a surface to the ligand  
18 cavity for interaction between at least one amino acid and the bound ligand. The C-terminal  
19 17 amino acids of the TR, referred to as the activation helix or AF-2 (an example of an LBD  
20 activation domain), are implicated in mediating hormone-dependent transcriptional activation.  
21 Although, mutations of key residues within the domain decrease ligand-dependent activation  
22 it was unclear until the present invention whether such mutations directly affected ligand  
23 coordination. Although some mutations of this domain have been noted to reduce or abolish  
24 ligand binding, other mutations in more distant sites of the LBD have a similar effect.

25 Activation domains among nuclear receptors display an analogous three dimensional  
26 relationship to the binding cavity, which is a region of the LBD that binds the molecular  
27 recognition domain of a ligand, i.e. the activation domain presents a portion of itself to the  
28 binding cavity (but necessarily the molecular recognition domain of the ligand). Many  
29 nuclear receptors are expected to have such domains, including the retinoid receptors, RAR  
30 and RXR, the glucocorticoid receptor GR, and the estrogen receptor ER. Based upon the  
31 TR's sequence, the domain is proposed to adopt an amphipathic helical structure.  $\beta$ -sheet or

1 mixed secondary structures, could be present as activation domains in less related nuclear  
2 receptors.

3 Within the activation domain, the highly conserved motif  $\Phi\Phi X E \Phi\Phi$ , where  $\Phi$   
4 represents a hydrophobic residue, is proposed to mediate interactions between the receptors  
5 and transcriptional coactivators. Several proteins have been identified which bind the TR in  
6 a hormone-dependent fashion. One of these, Trip1, is related to a putative yeast coactivator  
7 Sug1, and also interacts with both the C-terminal activation domain and a subset of the basal  
8 transcriptional machinery, suggesting a role in transactivation by the TR. Other proteins,  
9 such as RIP140, SRC1, (Oate, S.A. et. al., *Science* 270:1354-1357 (1995)) and TF-1 (see  
10 also Ledouarin, B., et. al., *EMBO J.* 14:2020-2033 (1995)), and GRIP-1 (Heery, E., et al.,  
11 *Nature* 387:733-736 (1997)) also interact with other nuclear receptors in a ligand dependent  
12 manner through the C-terminal domain. Binding of these proteins can be modulated using  
13 the TR ligands described herein especially those TR ligands with extensions that sterically  
14 hinder the interaction between the highly conserved motif and other proteins.

15 The C-terminal activation domain of the TR forms an amphipathic helix, H12, which  
16 nestles loosely against the receptor to form part of the hormone binding cavity. The helix  
17 packs with the hydrophobic residues facing inward towards the hormone binding cavity, and  
18 the charged residues, including the highly-conserved glutamate, extending into the solvent, as  
19 shown in FIG. 8. The activation helix of TR LBD presents Phe 401 to the ligand binding  
20 cavity and permits direct coordination with the hormone i.e. such amino acids interact with  
21 the ligand forming a van der waals contact with the plane of the outer phenyl ring. Phe 405  
22 also interacts with His 381, perhaps stabilizing its hydrogen bonding conformation, i.e. a  
23 favorable hydrogen bond interaction. Participation of Phe 401 and Phe 405 in binding  
24 hormone explains how mutation of these residues decreases hormone binding affinity.  
25 Furthermore, the impact of these mutations on activation likely derives from a role in  
26 stabilizing the domain in the bound structure through increased hydrogen bond interaction of  
27 dipole interactions. Glu 403 extends into the solvent, emphasizing its critical role in  
28 transactivation. In its observed conformation, presented on the surface as an ordered  
29 residue, against a background of predominantly hydrophobic surface, Glu 403 is available to  
30 interact with activator proteins described herein, as shown in FIG. 9. The other charged  
31 residues, Glu 405 and Asp 406 are disordered, as the helix frays at Phe 405.



Two other sequences in the TR,  $\tau 2$  and  $\tau 3$ , activate transcription when expressed as fusion proteins with a DNA-binding domain. The sequences, discovered in the TRB, correspond to TR- $\alpha$  residues Pro158-Ile168 in H1 ( $\tau 2$ ), and Gly290-Leu319 in H8 and H9 ( $\tau 3$ ). Unlike the C-terminal activation domain,  $\tau 2$  and  $\tau 3$  do not appear to represent modular structural units in the rat TR- $\alpha$  LBD, nor present a surface for protein-protein interactions: the critical aspartate/glutamate residues of  $\tau 3$  are located on two separate helices, and do not form a single surface; the charged residues of  $\tau 2$  are engaged in ion pair interactions with residues of the LBD. Thus,  $\tau 2$  and  $\tau 3$  may not function as activation domains in the context of the entire receptor.

#### Computational Methods For Designing A Nuclear Receptor LBD LIGAND

The elucidation of the three dimensional structure of a nuclear receptor ligand binding domain provides an important and useful approach for designing ligands to nuclear receptors using the computational methods described herein. By inspecting the **FIGURES** it can be determined that the nuclear receptor ligand is bound in a water inaccessible binding cavity in the LBD and that chemical moieties can be added to selected positions on the ligand. Such chemical modifications, usually extensions, can fill up the binding cavity represented in the **FIGURES** for a tighter fit (or less water) or can be used to disrupt or make contacts with amino acids not in contact with the ligand before the chemical modification was introduced or represented in a figure of the three dimensional model of the LBD. Ligands that interact with nuclear superfamily members can act as agonists, antagonists and partial agonists based on what ligand-induced conformational changes take place.

Agonists induce changes in receptors that place them in an active conformation that allows them to influence transcription, either positively or negatively. There may be several different ligand-induced changes in the receptor's conformation.

Antagonists, bind to receptors, but fail to induce conformational changes that alter the receptor's transcriptional regulatory properties or physiologically relevant conformations. Binding of an antagonist can also block the binding and therefore the actions of an agonist.

Partial agonists bind to receptors and induce only part of the changes in the receptors that are induced by agonists. The differences can be qualitative or quantitative. Thus, a

1 partial agonist may induce some of the conformation changes induced by agonists, but not  
2 others, or it may only induce certain changes to a limited extent.

#### 4 **Ligand-induced Conformational Changes**

5 As described herein, the unliganded receptor is in a configuration that is either  
6 inactive, has some activity or has repressor activity. Binding of agonist ligands induces  
7 conformational changes in the receptor such that the receptor becomes more active, either to  
8 stimulate or repress the expression of genes. The receptors may also have non-genomic  
9 actions. Some of the known types of changes and/or the sequelae of these are listed herein.

#### 11 ***Heat Shock Protein Binding***

12 For many of the nuclear receptors ligand binding induces a dissociation of heat shock  
13 proteins such that the receptors can form dimers in most cases, after which the receptors  
14 bind to DNA and regulate transcription.

15 Nuclear receptors usually have heat shock protein binding domains that present a  
16 region for binding to the LBD and can be modulated by the binding of a ligand to the LBD.  
17 Consequently, an extended chemical moiety (or more) from the ligand that stabilizes the  
18 binding or contact of the heat shock protein binding domain with the LBD can be designed  
19 using the computational methods described herein to produce a partial agonist or antagonist.  
20 Typically such extended chemical moieties will extend past and away from the molecular  
21 recognition domain on the ligand and usually past the buried binding cavity of the ligand.

#### 23 ***Dimerization and Heterodimerization***

24 With the receptors that are associated with the hsp in the absence of the ligand,  
25 dissociation of the hsp results in dimerization of the receptors. Dimerization is due to  
26 receptor domains in both the DBD and the LBD. Although the main stimulus for  
27 dimerization is dissociation of the hsp, the ligand-induced conformational changes in the  
28 receptors may have an additional facilitative influence. With the receptors that are not  
29 associated with hsp in the absence of the ligand, particularly with the TR, ligand binding can  
30 affect the pattern of dimerization/heterodimerization. The influence depends on the DNA  
31 binding site context, and may also depend on the promoter context with respect to other

1 proteins that may interact with the receptors. A common pattern is to discourage monomer  
2 formation, with a resulting preference for heterodimer formation over dimer formation on  
3 DNA.

4 Nuclear receptor LBDs usually have dimerization domains that present a region for  
5 binding to another nuclear receptor and can be modulated by the binding of a ligand to the  
6 LBD. Consequently, an extended chemical moiety (or more) from the ligand that disrupts  
7 the binding or contact of the dimerization domain can be designed using the computational  
8 methods described herein to produce a partial agonist or antagonist. Typically such extended  
9 chemical moieties will extend past and away from the molecular recognition domain on the  
10 ligand and usually past the buried binding cavity of the ligand.

#### 12 *DNA Binding*

13 In nuclear receptors that bind to hsp, the ligand-induced dissociation of hsp with  
14 consequent dimer formation allows, and therefore, promotes DNA binding. With receptors  
15 that are not associated (as in the absence of ligand), ligand binding tends to stimulate DNA  
16 binding of heterodimers and dimers, and to discourage monomer binding to DNA.

17 However, ligand binding to TR, for example, tends to decrease dimer binding on certain  
18 DNA elements and has minimal to no effect on increasing heterodimer binding. With DNA  
19 containing only a single half site, the ligand tends to stimulate the receptor's binding to  
20 DNA. The effects are modest and depend on the nature of the DNA site and probably on the  
21 presence of other proteins that may interact with the receptors. Nuclear receptors usually  
22 have DBDs that present a region for binding to DNA and this binding can be modulated by  
23 the binding of a ligand to the LBD. Consequently, an extended chemical moiety (or more)  
24 from the ligand that disrupts the binding or contact of the DBD can be designed using the  
25 computational methods described herein to produce a partial agonist or antagonist. Typically  
26 such extended chemical moieties will extend past and away from the molecular recognition  
27 domain on the ligand and usually past the buried binding cavity of the ligand.

#### 29 *Repressor Binding*

30 Receptors that are not associated with hsp in the absence of ligand frequently act as  
31 transcriptional repressors in the absence of the ligand. This appears to be due, in part, to

1 transcriptional repressor proteins that bind to the LBD of the receptors. Agonist binding  
2 induces a dissociation of these proteins from the receptors. This relieves the inhibition of  
3 transcription and allows the transcriptional transactivation functions of the receptors to  
4 become manifest.

#### 6 *Transcriptional Transactivation Functions*

7 Ligand binding induces transcriptional activation functions in two basic ways. The  
8 first is through dissociation of the hsp from receptors. This dissociation, with consequent  
9 dimerization of the receptors and their binding to DNA or other proteins in the nuclear  
10 chromatin allows transcriptional regulatory properties of the receptors to be manifest. This  
11 may be especially true of such functions on the amino terminus of the receptors.

12 The second way is to alter the receptor to interact with other proteins involved in  
13 transcription. These could be proteins that interact directly or indirectly with elements of the  
14 proximal promoter or proteins of the proximal promoter. Alternatively, the interactions  
15 could be through other transcription factors that themselves interact directly or indirectly with  
16 proteins of the proximal promoter. Several different proteins have been described that bind  
17 to the receptors in a ligand-dependent manner. In addition, it is possible that in some cases,  
18 the ligand-induced conformational changes do not affect the binding of other proteins to the  
19 receptor, but do affect their abilities to regulate transcription.

20 Nuclear receptors or nuclear receptor LBDs usually have activation domains  
21 modulated in part by a co-activator/co-repressor system that coordinately functions to present  
22 a region for binding to DNA, and can be modulated by the binding of a ligand to the LBD.  
23 Consequently, an extended chemical moiety (or more) from the ligand that disrupts the  
24 binding or contact of the activation domain with co-activator and/or co-repressor can be  
25 designed using the computational methods described herein to produce a partial agonist or  
26 antagonist. For instance, an agonist can be designed and/or selected which (1) blocks  
27 binding and/or dissociates co-repressor, and/or (2) promotes binding and/or association of a  
28 co-activator. An antagonist can be designed which (1) promotes binding and/or association  
29 of co-repressor, and/or (2) promotes binding and/or association of co-activator. Ratios of  
30 agonists and antagonists may be used to modulate transcription of the gene of interest.  
31 Selection can be accomplished in binding assays that screen for ligands having the desired

agonist or antagonist properties, including such ligands which induce conformational changes as described below. Suitable assays for such screening are described herein and in Shibata, H., *et al.* (*Recent Prog. Horm. Res.* 52:141-164 (1997)); Tagami, T., *et al.* (*Mol. Cell Biol.* 17(5):2642-2648 (1997)); Zhu, XG., *et al.* (*J. Biol. Chem.* 272(14):9048-9054 (1997)); Lin, B.C., *et al.* (*Mol. Cell Biol.* 17(10):6131-6138 (1997)); Kakizawa, T., *et al.* (*J. Biol. Chem.* 272(38):23799-23804 (1997)); and Chang, K. H., *et al.* (*Proc. Natl. Acad. Sci. USA* 94(17):9040-9045 (1997)). Typically such extended chemical moieties will extend past and away from the molecular recognition domain on the ligand and usually past the buried binding cavity of the ligand and in the direction of the activation domain, which is often a helix as seen in the three dimensional model shown in the **FIGURES** in two dimensions on paper or more conveniently on a computer screen.

### **Ligand-Induced Conformational Change**

Plasma proteins bind hormones without undergoing a conformational change through a static binding pocket formed between monomers or domains. For example, the tetrameric thyroid-binding plasma protein transthyretin forms a solvent-accessible hormone-binding channel at the oligomer interface. The structure of the protein is unchanged upon binding hormone with respect to the appearance of a buried binding cavity with a ligand bound.

However, the structural role for a ligand bound to a nuclear receptor LBD, like rat TR- $\alpha$  LBD, predicts that the receptor would differ in the bound and unbound states. In the absence of hormone, the receptor would possess a cavity at its core, uncharacteristic of a globular protein. A ligand (e.g. hormone) completes the hydrophobic core of the active receptor after it binds to the nuclear receptor. Ligand binding by the receptor is a dynamic process, which regulates receptor function by inducing an altered conformation.

An exact description of the hormone-induced conformational changes requires comparison of the structures of the liganded and the unliganded TR. The structure of the unliganded human RXR $\alpha$  may substitute as a model for the unliganded TR. The rat TR- $\alpha$  LBD and human RXR $\alpha$  LBDs adopt a similar fold, and it is likely that the structural similarity extends to the conformational changes after ligand binding.

There are three major differences between the two structures, which indeed appear to be the result of ligand binding. First, the bound rat TR- $\alpha$  LBD structure is more compact,

1 with the hormone tightly packed within the hydrophobic core of the receptor. By contrast,  
2 the unliganded human RXR $\alpha$  LBD contains several internal hydrophobic cavities. The  
3 presence of such cavities is unusual in folded proteins, and is likely a reflection of the  
4 unliganded state of the receptor. Two of these cavities were proposed as possible binding  
5 sites for 9-cis retinoic acid, though these multiple sites only partly overlap with the single  
6 buried binding cavity observed in the liganded rat TR- $\alpha$  LBD.

7 The second difference involves H11 in the rat TR- $\alpha$  LBD, which contributes part of  
8 the hormone binding cavity. H11, continuous in the rat TR- $\alpha$  LBD, is broken at Cys 432 in  
9 the RXR, forming a loop between H10 and H11 in the hRXR $\alpha$ . This residue corresponds to  
10 His381 in the TR, which provides a hydrogen bond to the outer ring hydroxyl of the ligand.  
11 Furthermore, the hormone binding cavity occupied by ligand in the rat TR- $\alpha$  LBD is  
12 interrupted in the hRXR $\alpha$  by the same loop, forming an isolated hydrophobic pocket in the  
13 RXR with H6 and H7. In the bound rat TR- $\alpha$  LBD, the corresponding helices H7 and H8  
14 are contiguous with the binding pocket, and enclose the hormone binding cavity from below.

15 The third difference between the two receptors is the position of the C-terminal  
16 activation domain. While the C-terminal activation domain forms  $\alpha$ -helices in both  
17 receptors, the domain in the rat TR- $\alpha$  LBD follows a proline-rich turn, and lies against the  
18 receptor to contribute part of the binding cavity. In contrast, the activation domain in the  
19 unliganded hRXR $\alpha$ , is part of a longer helix which projects into the solvent.

20 These differences lead to a model for an alternate conformation of the TR LBD  
21 assumed in the absence of ligand. In the unliganded TR, the subdomain of the receptor  
22 surrounding the hormone binding cavity is loosely packed, with the binding cavity occluded  
23 by a partly unstructured H11 providing a partial core for the receptor.

24 Upon binding hormone, residues which form a coil in the unbound receptor engage  
25 the ligand, and continues H11. The ordering of H11 could unblock the hydrophobic cavity,  
26 allowing H7 and H8 to interact with hormone. The extended hydrophobic cavity then  
27 collapses around the hormone, generating the compact bound structure.

28 It is possible to predict ligand-induced conformational changes in the C-terminal  
29 activation domain that rely, in part, on an extended structure in the unliganded TR that  
30 repacks upon ligand binding. The ligand-induced conformation change can be subtle since  
31 the amino acid sequence of the rat TR- $\alpha$  in the turn (393-PTELFPP-399) significantly

1 reduces the propensity of the peptide chain of the rat TR- $\alpha$  to form an  $\alpha$ -helix and therefore  
2 repacking can be accomplished with a minor change in volume.

3 After the ligand-induced conformational change occurs, it is likely that the  
4 conformation of the C-terminal activation domain in the bound structure changes packing  
5 compared to the unbound form of the receptor. Binding of the ligand improves the stability  
6 of the activation domain. The activation domain packs loosely even in the bound structure,  
7 as measured by the distribution of packing interactions for the entire LBD. The packing  
8 density for the activation domain, defined as the number of atoms within 4.5Å, is 1.5  
9 standard deviations below the mean. For comparison, another surface helix, H1, is 0.5  
10 standard deviations below the mean and the most poorly packed part of the structure, the  
11 irregular coil from residues Ile196-Asp206, is 2.0 standard deviations below the mean.  
12 Moreover, the majority of packing contacts for the C-terminal domain in the bound receptor  
13 are provided either by residues which interact with ligand, such as His381, or by the ligand  
14 itself. The conformation of these residues can be expected to be different in the bound and -  
15 unbound receptors, and by extension the conformation of C-terminal activation domain which  
16 relies upon these interactions. Without the stabilization provided by a bound ligand, it is  
17 likely that the C-terminal activation domain is disordered prior to hormone binding.

18 The interrelation of ligand-induced conformational changes is evident as described  
19 herein. For example, His381 from H11 and Phe405 from H12 interact in the bound  
20 structure to provide a specific hydrogen bond to the phenolic hydroxyl. The ligand-induced  
21 changes which affect H11 and H12 are reinforcing, and lead to the formation of the compact,  
22 bound state.

23 Comparison of the TR- $\alpha$  and TR- $\beta$  LBD structures shows similar packing of the  
24 helices when complexed with the ligand Triac.

25

## 26 COMPUTATIONAL METHODS USING THREE DIMENSIONAL MODELS AND EXTENSIONS OF 27 LIGANDS

28 The three-dimensional structure of the liganded TR receptor is unprecedented, and  
29 will greatly aid in the development of new nuclear receptor synthetic ligands, such as thyroid  
30 receptor antagonists and improved agonists, especially those that bind selectively to one of  
31 the two TR isoforms ( $\alpha$  or  $\beta$ ). In addition, this receptor superfamily is overall well suited to

1 modern methods including three-dimensional structure elucidation and combinatorial  
2 chemistry such as those disclosed in EP 335 628, U.S. patent 5,463,564, which are  
3 incorporated herein by reference. Structure determination using X-ray crystallography is  
4 possible because of the solubility properties of the receptors. Computer programs that use  
5 crystallography data when practicing the present invention will enable the rational design of  
6 ligand to these receptors. Programs such as RASMOL can be used with the atomic  
7 coordinates from crystals generated by practicing the invention or used to practice the  
8 invention by generating three dimensional models and/or determining the structures involved  
9 in ligand binding. Computer programs such as INSIGHT and GRASP allow for further  
10 manipulation and the ability to introduce new structures. In addition, high throughput  
11 binding and bioactivity assays can be devised using purified recombinant protein and modern  
12 reporter gene transcription assays described herein and known in the art in order to refine the  
13 activity of a CDL.

14 Generally the computational method of designing a nuclear receptor synthetic ligand -  
15 comprises two steps:

16 1) determining which amino acid or amino acids of a nuclear receptor LBD interacts  
17 with a first chemical moiety (at least one) of the ligand using a three dimensional model of a  
18 crystallized protein comprising a nuclear receptor LBD with a bound ligand, and

19 2) selecting a chemical modification (at least one) of the first chemical moiety to  
20 produce a second chemical moiety with a structure to either decrease or increase an  
21 interaction between the interacting amino acid and the second chemical moiety compared to  
22 the interaction between the interacting amino acid and the first chemical moiety.

23 As shown herein, interacting amino acids form contacts with the ligand and the center  
24 of the atoms of the interacting amino acids are usually 2 to 4 angstroms away from the center  
25 of the atoms of the ligand. Generally these distances are determined by computer as  
26 discussed herein and in McRee 1993, however distances can be determined manually once  
27 the three dimensional model is made. Examples of interacting amino acids are described in  
28 Appendix 2. See also Wagner *et al.*, *Nature* 378(6558):670-697 (1995) for stereochemical  
29 figures of three dimensional models. More commonly, the atoms of the ligand and the atoms  
30 of interacting amino acids are 3 to 4 angstroms apart. The invention can be practiced by  
31 repeating steps 1 and 2 to refine the fit of the ligand to the LBD and to determine a better



1 ligand, such as an agonist. As shown in the FIGURES the three dimensional model of TR  
2 can be represented in two dimensions to determine which amino acids contact the ligand and  
3 to select a position on the ligand for chemical modification and changing the interaction with  
4 a particular amino acid compared to that before chemical modification. Structural  
5 comparison of LBD isoforms complexed with the same or similar ligand permit identification  
6 of fiducial and adjustable amino acids that can be exploited in designing isoform-specific  
7 ligands through chemical modification. "Fiducial" refers to amino acids that form rigid  
8 features of the ligand binding cavity. "Adjustable" refers to amino acids that form less rigid  
9 features of the ligand binding cavity. The chemical modification may be made using a  
10 computer, manually using a two dimensional representation of the three dimensional model  
11 or by chemically synthesizing the ligand. The three dimensional model may be made using  
12 Appendix 2 and the FIGURES. As an additional step, the three dimensional model may be  
13 made using atomic coordinates of nuclear receptor LBDs from crystallized protein as known  
14 in the art, see McRee 1993 referenced herein.

15 The ligand can also interact with distant amino acids after chemical modification of  
16 the ligand to create a new ligand. Distant amino acids are generally not in contact with the  
17 ligand before chemical modification. A chemical modification can change the structure of  
18 the ligand to make as new ligand that interacts with a distant amino acid usually at least 4.5  
19 angstroms away from the ligand. Often distant amino acids will not line the surface of the  
20 binding cavity for the ligand, as they are too far away from the ligand to be part of a pocket  
21 or surface of the binding cavity.

22 The interaction between an atom of a LBD amino acid and an atom of an LBD ligand  
23 can be made by any force or attraction described in nature. Usually the interaction between  
24 the atom of the amino acid and the ligand will be the result of a hydrogen bonding  
25 interaction, charge interaction, hydrophobic interaction, van der waals interaction or dipole  
26 interaction. In the case of the hydrophobic interaction it is recognized that this is not a per  
27 se interaction between the amino acid and ligand, but rather the usual result, in part, of the  
28 repulsion of water or other hydrophilic group from a hydrophobic surface. Reduction or  
29 enhancement of the interaction of the LBD and a ligand can be measured by standard binding  
30 procedures, calculating or testing binding energies, computationally or using thermodynamic  
31 or kinetic methods as known in the art.

1 Chemical modifications will often enhance or reduce interactions of an atom of a LBD  
2 amino acid and an atom of an LBD ligand. Steric hinderance will be a common means of  
3 changing the interaction of the LBD binding cavity with the activation domain. Chemical  
4 modifications are preferably introduced at C-H, C- and C-OH position in ligands, where the  
5 carbon is part of the ligand structure which remains the same after modification is complete.  
6 In the case of C-H, C could have 1, 2 or 3 hydrogens, but usually only one hydrogen will be  
7 replaced. The H or OH are removed after modification is complete and replaced with the  
8 desired chemical moiety.

9 Because the thyroid receptor is a member of the larger superfamily of hormone-  
10 binding nuclear receptors, the rules for agonist and antagonist development will be  
11 recognized by one skilled in the art as useful in designing ligands to the entire superfamily.  
12 Examining the structures of known agonists and antagonists of the estrogen and androgen  
13 receptors supports the generality of antagonist mechanism of action as shown in FIG. 10.

14 The overall folding of the receptor based on a comparison of the reported structure of  
15 the unliganded RXR and with amino acid sequences of other superfamily members reveals  
16 that the overall folding of receptors of the superfamily is similar. Thus, it is predicted from  
17 the structure that there is a general pattern of folding of the nuclear receptor around the  
18 agonist or antagonist ligand.

19 The three dimensional structure of a nuclear receptor with a ligand bound leads to the  
20 nonobvious observation that a nuclear receptor folds around agonist ligands, as the binding  
21 cavity fits the agonist, especially the agonist's molecular recognition domain, and antagonists  
22 commonly have chemical structures that extend beyond the ligand, especially the agonist, and  
23 would prohibit folding of the receptor around the ligand to form a buried binding cavity or  
24 other groups that have the same effect. The location of the extension could affect the folding  
25 in various ways as indicated by the structure. Such extensions on antagonists are shown in  
26 FIG. 10 for various receptors and compared to the corresponding agonist.

27 For example, an extension towards the carboxy-terminal activation helix affects the  
28 packing/folding of this helix into the body of the receptor. This in turn can affect the ability  
29 of this portion of the nuclear receptor to interact with other proteins or other portions of the  
30 receptor, including transcriptional transactivation functions on the opposite end of the linear  
31 receptor, or the receptor's amino terminus that may interact directly or indirectly with the

1 carboxy-terminal transactivation domain (including helix 12). Extensions in this direction  
2 can also affect the packing of helix 11 of TR (or its analogous helix in nuclear receptors) into  
3 the body of the receptor and selectively affect dimerization and heterodimerization of  
4 receptors. An extension pointing towards helix 1 can affect the relationship of the DNA  
5 binding domain and hinge regions of the receptors with the ligand binding domain and  
6 selectively or in addition affect the receptors' binding to DNA and/or interactions of  
7 receptors with proteins that interact with this region of the receptor. Other extensions  
8 towards helix 11 can be made to affect the packing of this helix and helices 1 and 10 and  
9 thereby homo- and hetero-dimerization. Such chemical modifications can be assessed using  
10 the computational methods described herein. It is also possible that, in some cases,  
11 extensions may protrude through the receptor that is otherwise completely or incompletely  
12 folded around the ligand. Such protruding extensions could present a steric blockade to  
13 interactions with co-activators or other proteins.

14 The three dimensional structure with the ligand buried in the binding cavity  
15 immediately offers a simple description of a nuclear receptor that has a binding cavity that  
16 contains hinges and a lid, composed of one or more structural elements, that move to  
17 accommodate and surround the ligand. The ligand to TR can be modified on specific sites  
18 with specific classes of chemical groups that will serve to leave the lid and hinge region in  
19 open, partially open or closed states to achieve partial agonist or antagonist functions. In  
20 these states, the biological response of the TR is different and so the structure can be used to  
21 design particular compounds with desired effects.

22 Knowledge of the three-dimensional structure of the TR-T<sub>3</sub> complex leads to a general  
23 model for agonist and antagonist design. An important novel feature of the structural data is  
24 the fact that the T<sub>3</sub> ligand is completely buried within the central hydrophobic core of the  
25 protein. Other ligand-receptor complexes belonging to the nuclear receptor superfamily will  
26 have a similarly buried ligand binding site and therefore this model will be useful for  
27 agonist/antagonist design for the entire superfamily.

28 When design of an antagonist is desired, one needs either to preserve the important  
29 binding contacts of natural hormone agonist while incorporating an "extension group" that  
30 interferes with the normal operation of the ligand-receptor complex or to generate the  
31 requisite binding affinity through the interactions of the extensions with receptor domains.

1       The model applied to antagonist design and described herein is called the "Extension  
2 Model." Antagonist compounds for nuclear receptors should contain the same or similar  
3 groups that facilitate high-affinity binding to the receptor, and in addition, such compounds  
4 should contain a side chain which may be large and/or polar. This side chain could be an  
5 actual extension, giving it bulk, or it could be a side group with a charge function that differs  
6 from the agonist ligand. For example, substitution of a CH<sub>3</sub> for CH<sub>2</sub>OH at the 21-position,  
7 and alteration at the 11-position from an OH group to a keto group of cortisol generates  
8 glucocorticoid antagonist activity (Robsseau, G.G., *et. al.*, *J. Mol. Biol.* 67:99-115 (1972)).  
9 However, in most cases effective antagonists have more bulky extensions. Thus, the  
10 antiglucocorticoid (and antiprogesterin) RU486 contains a bulky side group at the 11-position  
11 (Horwitz, K.B. *Endocrine Rev.* 13:146-163 (1992)). The antagonist compound will then  
12 bind within the buried ligand binding site of the receptor with reasonably high affinity (100  
13 nM), but the extension function will prevent the receptor-ligand complex from adopting the  
14 necessary conformation needed for transcription factor function. The antagonism (which  
15 could be in an agonist or antagonist) may manifest itself at the molecular level in a number  
16 of ways, including by preventing receptor homo/heterodimer formation at the HRE, by  
17 preventing coactivator binding to receptor monomers, homodimers or homo/heterodimers, or  
18 by a combination of these effects which otherwise prevent transcription of hormone  
19 responsive genes mediated by ligand-induced effects on the HRE. There are several  
20 antagonist compounds for nuclear receptors in the prior art (see also Horwitz, K.B.,  
21 *Endocrine Rev.* 13:146-163 (1992), Raunnaud J.P. *et. al.*, *J. Steroid Biochem.* 25:811-833  
22 (1986), Keiel S., *et. al.*, *Mol. Cell. Biol.* 14:287-298 (1994) whose antagonist function can  
23 be explained by the extension hypothesis. These compounds are shown in FIG. 10 along  
24 with their agonist counterparts. Each of these antagonists contains a large extension group  
25 attached to an agonist or agonist analogue core structure. Importantly, these antagonist  
26 compounds were discovered by chance and not designed with a structure-function hypothesis  
27 such as the extension principle.

28       One method of design of a thyroid antagonist using the extension hypothesis is  
29 provided below as a teaching example. The three-dimensional structure of the TR- $\alpha$  Dimit  
30 complex combined with structure-activity data published in the prior art, especially those  
31 reference herein, can be used to establish the following ligand-receptor interactions which are

1 most critical for high-affinity ligand binding. A physical picture of these interactions is  
2 shown in FIG. 6. The figure describes the isolated essential contacts for ligand binding.  
3 Because the ligand is buried in the center of the receptor, the structural spacing between  
4 these isolated interactions is also important. Thus, our present knowledge of this system  
5 dictates that, for this example, a newly designed ligand for the receptor must contain a  
6 thyronine structural skeleton, or two substituted aryl groups joined by a one-atom spacer.

7 The general structure for an antagonist designed by the extension hypothesis is  
8 exemplified in the following general description of the substituents of a TR antagonist  
9 (referring to Formula 1): R1 can have anionic groups such as a carboxylate, phosphonate,  
10 phosphate, sulfate or sulfite and is connected to the ring with a 0 to 3 atom linker,  
11 comprising one or more C, O, N, S atoms, and preferably a 2 carbon linker. Such R1 can  
12 be optionally substituted with an amine (e.g. -NH<sub>2</sub>). R3 and R5 are small hydrophobic  
13 groups such as -Br, -I, or -CH<sub>3</sub>. R3 and R5 can be the same substituents or different. R<sub>3</sub>'  
14 can be a hydrophobic group that may be larger than those of R3 and R5, such as -I, -CH<sub>3</sub>, -  
15 isopropyl, -phenyl, -benzyl, 5 and 6 ring heterocycles. R<sub>4</sub>' is a group that can participate in  
16 a hydrogen bond as either a donor or acceptor. Such groups include -OH, -NH<sub>2</sub>, and -SH.  
17 R<sub>5</sub>' is an important extension group that makes this compound an antagonist. R<sub>5</sub>' can be a  
18 long chain alkyl (e.g. 1 to 9 carbons, straight chain or branched), aryl (benzyl, phenyl and  
19 substituted benzyl and phenyl rings (e.g. with halogen, alkyl (1 and 5 carbons) and optionally  
20 connected to the ring by a -CH<sub>2</sub>-), heterocycle (e.g. 5 or 6 atoms, preferably 5 carbons and  
21 1 nitrogen, or five carbons), which can optionally include polar (e.g. -OH, -NH<sub>2</sub>, and -SH),  
22 cationic (e.g. -NH<sub>3</sub>, N(CH<sub>3</sub>)<sub>3</sub>), or anionic (carboxylate, phosphonate, phosphate or sulfate)  
23 groups. R<sub>5</sub>' can also be a polar (e.g. -OH, -NH<sub>2</sub>, and -SH), cationic (e.g. -NH<sub>3</sub>, -  
24 N(CH<sub>3</sub>)<sub>3</sub>), and anionic (carboxylate, phosphonate, phosphate or sulfate) groups. X is the  
25 spacer group that appropriately positions the two aromatic rings. This group is usually a  
26 one-atom spacer, such as O, S, SO, SO<sub>2</sub>, NH, NZ where Z is an alkyl, CH<sub>2</sub>, CHOH, CO,  
27 C(CH<sub>3</sub>)OH, and C(CH<sub>3</sub>)(CH<sub>3</sub>). X also may be NR<sub>7</sub>, CHR<sub>7</sub>, CR<sub>7</sub>, R<sub>7</sub>, where R<sub>7</sub> is an alkyl,  
28 aryl or 5- or 6-membered heterocyclic aromatic. R2, R6, R2' and R6' can be -F, and -Cl  
29 and are preferably H.

30 A TR ligand can also be described as a substituted phenylated 3,5 diiodo tyrosine with  
31 substituted R5' and R3' groups. R5' can be a long chain alkyl (e.g. 4 to 9 carbons, straight

1 chain or branched), aryl (benzyl, phenyl and substituted benzyl and phenyl rings (e.g. with  
2 halogen, alkyl (1 and 5 carbons) and optionally connected to the ring by a -CH<sub>2</sub>-),  
3 heterocycle (e.g. 5 or 6 atoms, preferably 5 carbons and 1 nitrogen, or five carbons), which  
4 can optionally include polar (e.g. -OH, -NH<sub>2</sub>, and -SH), cationic (e.g. -NH<sub>3</sub>, N(CH<sub>3</sub>)<sub>3</sub>), or  
5 anionic (carboxylate, phosphonate, phosphate or sulfate) groups. R<sub>5</sub>' can also be a polar  
6 (e.g. -OH, -NH<sub>2</sub>, and -SH), cationic (e.g. -NH<sub>3</sub>, N(CH<sub>3</sub>)<sub>3</sub>), and anionic (carboxylate,  
7 phosphonate, phosphate or sulfate) groups. R<sub>3</sub>' can be -IsoPr, halogen, -CH<sub>3</sub>, alkyl (1 to 6  
8 carbons) or aryl (benzyl, phenyl and substituted benzyl and phenyl rings (e.g. with halogen,  
9 alkyl (1 and 5 carbons) and optionally connected to the ring by a -CH<sub>2</sub>-), heterocycle (e.g. 5  
10 or 6 atoms, preferably 5 carbons and 1 nitrogen, or five carbons), which can optionally  
11 include polar (e.g. -OH, -NH<sub>2</sub>, and -SH), cationic (e.g. -NH<sub>3</sub>, N(CH<sub>3</sub>)<sub>3</sub>), or anionic  
12 (carboxylate, phosphonate, phosphate or sulfate) groups.

13 A TR antagonist can also be a modified T<sub>3</sub> agonist (having a biphenyl structure)  
14 wherein R<sub>5</sub>' is alkyl, aryl, 5- or 6-membered heterocyclic aromatic, heteroalkyl, heteroaryl,  
15 arylalkyl, heteroaryl alkyl, polyaromatic, polyheteroaromatic, polar or charged groups,  
16 wherein said R<sub>5</sub>' may be substituted with polar or charged groups. The R<sub>5</sub>' groups are  
17 defined, as described herein.

18 Using these methods the ligands of this example preferably have the following  
19 properties:

- 20 1. The compounds should bind to the TR with high affinity (for example 100 nM).
- 21 2. The compounds should bind the receptor in the same basic orientation as the
- 22 natural hormone.
- 23 3. The extension group R<sub>5</sub>' should project toward the activation helix (C-terminal
- 24 helix) of the receptor.
- 25 4. The appropriate substituent at R<sub>5</sub>' should perturb the activation helix from its
- 26 optimal local structure needed for mediating transcription.

27 Antagonists may also be designed with multiple extensions in order to block more  
28 than one aspect of the folding at any time.

29 TR ligands (e.g. super agonists) can be designed (and synthesized) to enhance the  
30 interaction of at least one amino acid with at least one chemical moiety on the ligand's  
31 molecular recognition domain. One method is to enhance the charge and polar interactions

1 by replacing the carboxylate of T<sub>3</sub> (R1 position) with phosphonate, phosphate, sulfate or  
2 sulfite. This enhances the interaction with Arg 262, Arg 266 and Arg 228. The interaction  
3 of at least one amino acid with at least one chemical moiety on the ligand's molecular  
4 recognition domain can also be enhanced by increasing the size of R1 group to fill the space  
5 occupied by water when Dimit is bound (referring to R1). Preferably the group has a  
6 complementary charge and hydrophobicity to the binding cavity.

7 Another way of improving the interaction of at least one amino acid with at least one  
8 chemical moiety on the ligand's molecular recognition domain is to restrict the conformation  
9 of the dihedral angle between the two phenyl rings of the thyronine ligand in solution. In  
10 solution the planes of two phenyl rings are orthogonal where the dihedral angle is 90°. In  
11 the TR Dimit structure, the dihedral angle is close to 60°. A TR ligand design that fixes the  
12 angle between the two phenyl rings will lead to tighter binding. Such a ligand may be made  
13 by connecting the R6' and the R5 positions of a thyronine or a substituted thyronine-like  
14 biphenyl. The size of the cyclic connection can fix the angle between the two phenyl rings.-  
15 Referring specifically to Formula 1, the following cyclic modifications are preferred: 1) R<sub>5</sub> is  
16 connected to R<sub>6</sub>', 2) R<sub>3</sub> is connected to R<sub>2</sub>' or 3) R<sub>5</sub> is connected to R<sub>6</sub>' and R<sub>3</sub> is connected  
17 to R<sub>2</sub>'. The connections can be made by an alkyl or heteroalkyl chain having between 1 to 6  
18 atoms and preferably from 2 to 4 carbon atoms or other atoms. Any position of the  
19 heteroalkyl chain can be N, O, P or S. The S and P heteroatoms along said heteroalkyl chain  
20 are in any of their possible oxidative states. The N heteroatom or any carbon along the alkyl  
21 or heteroalkyl chain may have one or more Z substituents, wherein Z is alkyl, heteroalkyl,  
22 aryl, heteroaryl, 5- or 6-membered heterocyclic aromatic. These compounds can be claimed  
23 with the proviso that Formula 1 does not include any prior art compound as of the priority  
24 filing date of this application.

25 The interaction of at least one amino acid with at least one chemical moiety on the  
26 ligand's molecular recognition domain can also be enhanced by selecting a chemical  
27 modification that fills the unfilled space between a TR ligand and the LBD in the area of the  
28 bridging oxygen (such as in T3, Triac or Dimit). Thus, a slighter larger moiety that replaces  
29 the ether oxygen can enhance binding. Such a linker may be a mono- or geminal-  
30 disubstituted carbon group. A group approximately the same size as oxygen but with greater

1 hydrophobicity is preferred as well as small, hydrophobic groups for the disubstituted  
2 carbon.

3       Compounds of Formula I or derivatives thereof that modulate TR activity also may be  
4 designed and selected to interact with a conformationally constrained structural feature of a  
5 TR LBD that is conserved among TR LBD isoforms to increase TR-specific selectivity.  
6 Conserved structural features of a TR LBD include residues found in equivalent positions of  
7 TR LBD isoforms which interact with a conserved structural feature of a compound  
8 comprising the biphenyl scaffold ( $\phi$ -X- $\phi$ ) or a single phenyl scaffold ( $\phi$ -X or X- $\phi$ ) of  
9 Formula I. Conformationally constrained structural features of a TR LBD include residues  
10 that have their natural flexible conformations fixed by various geometric and physical-  
11 chemical constraints, such as local backbone, local side chain, and topological constraints.  
12 These types of constraints are exploited to restrict positioning of atoms involved in receptor-  
13 ligand recognition and binding. For example, comparison of atomic models of TR LBD  
14 isoforms bound to thyronine and thyronine-like ligands reveal that certain residues which  
15 contact the ligands are restricted to particular topological shapes and angles of rotation about  
16 bonds. These include Met259, Leu276, Leu292, His381, Gly290, Ile221, and Phe401 of  
17 TR- $\alpha$ . The corresponding positions in TR- $\beta$  include Met313, Leu330, Leu346, His435,  
18 Gly344, Ile275 and Phe455, respectively.

19       Selectivity imparted by conformationally constrained features of both the receptor and  
20 compound are of particular interest. For example, compounds of Formula I comprising  
21 constrained cyclic carbons and substituent groups that interact with a constrained feature of a  
22 TR LBD can be exploited to further increase binding specificity while reducing the potential  
23 for cross-over interaction with other receptors. These include hydrophobic and/or  
24 hydrophilic contacts between constrained residues of a TR LBD and atomic groups of the  
25 following constituents of the compound in reference to Formula I: (i) the biphenyl rings; (ii)  
26 the R3-substituent; (iii) the R3'-substituent; and (iv) the R4'-substituent.

27       For example, contacts to the phenyl moiety comprising the R1, R2, R3, R5 and R6  
28 substituents, i.e., the ring proximal to the polar pocket (the "inner ring"), include a cycle  
29 carbon atom that interacts with an atom of a hydrophobic residue of a TR LBD, such as a  
30 carbon and oxygen atom of Met259 and a carbon atom of Leu276 of TR- $\alpha$ , or Met313 and  
31 Leu330 of TR- $\beta$ , where the cycle carbon is about 3.0 to 4.0Å from the atom of the



hydrophobic group. For example, comparison of TR- $\alpha$  complexed with T3 and TR- $\beta$  complexed with GC-1 reveals the following conserved inner ring contacts:

Ligand	TR LBD		
	TR- $\alpha$ Residue	Atom	Distance
T3/Atom			
C11	Met259	C	3.95
C11	Met259	O	3.59
C11	Met259	CB	3.77
C7	Leu276	CD2	3.80
C9	Leu276	CD2	3.70
GC1/Atom	TR- $\beta$ Residue	Atom	Distance
C11	Met313	C	3.85
C11	Met313	O	3.41
C11	Met313	CB	3.79
C7	Leu330	CD2	3.56
C9	Leu330	CD2	3.63

Contacts to the phenyl moiety comprising the R2', R3', R4', R5' and R6' substituents, i.e., the ring distal to the polar pocket (the "outer ring"), include a cyclic carbon atom that interacts with an atom of a hydrophobic residue of a TR LBD, such as a carbon atom of Leu292 of TR- $\alpha$ , or Leu346 of TR- $\beta$ , where the cyclic carbon atom is about 3.0 to 4.0A from the atom of the hydrophobic residue. For example, comparison of TR- $\alpha$  complexed with T3 and TR- $\beta$  complexed with GC-1 reveals the following conserved outer ring contacts:

Ligand	TR LBD		
	TR- $\alpha$ Residue	Atom	Distance
T3/Atom			
C6	Leu292	CD2	3.58
C8	Leu292	CD2	3.50
GC1/Atom	TR- $\beta$ Residue	Atom	Distance
C6	Leu346	CD2	3.77
C8	Leu346	CD2	3.80

Contacts to the R3-substituent include an atom that interacts with a carbon atom of a hydrophobic residue of a TR LBD, such as Ile221 of TR- $\alpha$ , or Ile275 of TR- $\beta$ , where the R3-substituent atom is about 3.0 to 4.0A from the carbon atom of the hydrophobic residue.

1 For example, comparison of TR- $\alpha$  complexed with T3 and TR- $\beta$  complexed with GC-1  
 2 reveals the following conserved R3'-substituent contacts:

Ligand	TR LBD		
	TR- $\alpha$ Residue	Atom	Distance
T3/Atom I1	Ile221	CG1	4.01
GC1/Atom C19	TR- $\beta$ Residue Ile275	Atom CG1	Distance 3.98

12 Contacts to the R3'-substituent include an atom that interacts with an atom of a  
 13 hydrophobic or hydrophilic residue of a TR LBD, such as an oxygen atom of Gly290 of TR-  
 14  $\alpha$ , or Gly344 of TR- $\beta$ , where the R3'-substituent atom is about 3.0 to 4.0A from the atom of  
 15 the hydrophobic or hydrophilic residue. For example, comparison of TR- $\alpha$  complexed with  
 16 T3 and TR- $\beta$  complexed with GC-1 reveals the following conserved R4'-substituent, phenolic  
 17 hydroxyl contacts:

Ligand	TR LBD		
	TR- $\alpha$ Residue	Atom	Distance
T3/Atom I2	Gly290	O	3.50
GC1/Atom C18	TR- $\beta$ Residue Gly344	Atom O	Distance 3.60

27 Contacts to the R4'-substituent comprising a phenolic hydroxyl include carbon and  
 28 oxygen atoms that interact with a hydrophobic or hydrophilic residue of a TR LBD, such as  
 29 a carbon and nitrogen atom of His381 of TR- $\alpha$ , or His435 of TR- $\beta$ , where the R4'-  
 30 substituent atom is about 2.0 to 4.0A from an atom of the hydrophobic or hydrophilic  
 31 residue. For example, comparison of TR- $\alpha$  complexed with T3 and TR- $\beta$  complexed with  
 32 GC-1 reveals the following conserved R4'-substituent, phenolic hydroxyl contacts:

Ligand	TR LBD		
T3/Atom	TR- $\alpha$ Residue	Atom	Distance
C10	His381	CD2	3.97
O1	His381	CD2	3.39
O1	His381	CE1	3.82
C8	His381	NE2	3.47
C10	His381	NE2	3.55
O1	His381	NE2	2.70

GC1/Atom	TR- $\beta$ Residue	Atom	Distance
C10	His435	CD2	3.89
O1	His435	CD2	3.64
O1	His435	CE1	3.79
C8	His435	NE2	3.44
C10	His435	NE2	3.33
O1	His435	NE2	2.77

Contacts to the R4'-substituent also may include an atom that interacts with a carbon atom of a hydrophobic residue of a TR LBD, such as Phe401 of TR- $\alpha$ , or Phe455 of TR- $\beta$ , for defining agonist activity, i.e., proper presentation of helix-12 (H12) of the TR LBD following ligand binding. The R4'-substituent atom is about 3.0 to 4.0Å from the carbon atom of the hydrophobic group. For example, comparison of TR- $\alpha$  complexed with T3 and TR- $\beta$  complexed with GC-1 reveals the following conserved R4'-substituent contacts:

Ligand	TR LBD		
T3/Atom	TR- $\alpha$ Residue	Atom	Distance
O1	Phe401	CE1	3.52
O1	Phe401	CZ	3.32
GC1/Atom	TR- $\beta$ Residue	Atom	Distance
O1	Phe455	CE1	3.40
O1	Phe455	CZ	3.22

Comparison of atomic models of TR LBD isoforms complexed with the same and/or different ligands therefore facilitates the identification of new compounds that fit spacially and preferentially into a TR LBD. Modeling, comparison of TR-ligand overlays, and comparison of TR LBD isoforms also permit identification of conformationally conserved structural features of TR LBD/ligand contacts. Exploiting conformational constraints of the LBD-ligand interaction identified by such methods therefore improves the design and

1 identification of new compounds having increased selectivity for binding a particular type of  
2 nuclear receptor, such as TR.

3

#### 4 **TR- $\alpha$ AND TR- $\beta$ SELECTIVITY FOR THE THYROID HORMONE RECEPTOR**

5 Using the method described herein ligands can be designed that selectively bind to the  
6 alpha more than the beta TR or vice versa. The X-ray crystallographic structure of the rat  
7 TR- $\alpha$  LBD provides insight into design of such ligands.

8 The three dimensional structure reveals that the major difference between the TR- $\alpha$   
9 and TR- $\beta$  in the ligand binding cavity resides in amino acid Ser 277 (with the side group  
10 -CH<sub>2</sub>OH) in the rat TR- $\alpha$  and whose corresponding residue is 331, asparagine (with the side  
11 group -CH<sub>2</sub>CONH<sub>2</sub>), in the human TR- $\beta$ . The side chain in human TR- $\beta$  is larger, charged  
12 and has a different hydrogen bonding potential, which would allow the synthesis of  
13 compounds that discriminate between this difference. The Ser277 (Asn331 in TR- $\beta$ ) forms  
14 part of the polar pocket of the TR LBD, indicating that for TR- $\alpha$  versus TR- $\beta$   
15 discrimination, ligands can be designed to contain chemical modification of the R1-  
16 substituent with reference to Formula I that exploit this difference.

17 For example, in the complex of TR- $\alpha$  with Triac, Ser277 does not participate in  
18 ligand binding. The absence of a role for Ser277 (Asn331 in beta) is consistent with the  
19 equal affinity of Triac for the alpha and beta isoforms, and indirectly supports the contention  
20 that alpha/beta selectivity resides in the amino acid substitution Ser277 to Asn331 and its  
21 interaction with Arg228. The effect of the amino acid substitution is further evident when  
22 the interactions of Asn331 and Arg282 in the structures of the TR- $\beta$  LBD complexed with  
23 GC-1 or Triac are compared with those of Ser277 and Arg228 in the TR- $\alpha$  LBD. In the  
24 complex with GC-1, Asn331 forms a hydrogen bond to Arg282, which in turn forms a  
25 hydrogen bond with the carboxylate of GC-1, a pattern that resembles the interactions of  
26 Ser277 and Arg228 in the complexes of the TR- $\alpha$  LBD complexed with T<sub>3</sub> or Triac.  
27 However, in the complex of TR- $\beta$  with Triac, Arg282 rotates away from Asn331 and the  
28 ligand, instead forming hydrogen bonds to residues Thr287 and Asp291 of H3. Therefore,  
29 differences exist between the two isoforms in the conformation of the polar pocket,  
30 depending on the nature of the ligand R1-substituent, indicating that certain substituents may  
31 interact preferentially with the conformation of a given isoform.

1 Comparing overlays of various ligands bound to the TR- $\alpha$  versus TR- $\beta$  LBDs shows  
2 the positioning of the ligand to be very similar. Surprisingly, comparison of the volume and  
3 area for the TR- $\alpha$  and TR- $\beta$  LBDs bound by the same or different ligands unexpectedly  
4 shows that the cubic space or volume available for accommodating ligand binding by the TR-  
5  $\beta$  LBD ( $645 \pm 28.28 \text{ \AA}^3$ ) is larger and more flexible than that of the TR- $\alpha$  LBD ( $596.25 \pm$   
6  $7.97 \text{ \AA}^3$ ) (Table 1). The volume of the ligand binding cavity for TR- $\alpha$  varies over a narrow  
7 range of about 8+, with a maximum difference of about 16+. In contrast, the volume of  
8 the ligand binding cavity for TR- $\beta$  differs by nearly 40+ between the complexes with GC-1  
9 and Triac. There also is a difference in the volume of the ligand binding cavity when  
10 comparing the same ligand bound to TR- $\alpha$  and TR- $\beta$ . For example, TR- $\alpha$  and TR- $\beta$   
11 complexed with Triac differ in LBD volume by about  $36 \text{ \AA}^3$ . Comparison of TR- $\alpha$  and TR-  
12  $\beta$  bound to Dimit and GC-1, respectively, which ligands have similar volume/area and  
13 superpositioned architecture, show that the difference in LBD volume is about  $75 \text{ \AA}^3$ . These  
14 differences are attributed primarily to variable movement and interaction of side chain groups  
15 with ligand substituents of the phenyl moiety ( $\phi$ ) of the biphenyl scaffold ( $\phi$ -X- $\phi$ ) located  
16 proximal to the polar pocket, e.g., R1-substituents in reference to Formula I. In contrast,  
17 the volume available in the hydrophobic pocket for both the TR- $\alpha$  and TR- $\beta$  LBDs is  
18 substantially the same. For example, binding of Triac to the TR- $\beta$  LBD displaces the side  
19 chain of Arg 282 providing approximately  $60 \text{ \AA}^3$  in the polar pocket cavity, exposing the  
20 polar pocket to bulk solvent exchange. For GC1 bound to the TR- $\beta$  LBD, approximately  $14$   
21  $\text{ \AA}^3$  is due to side chain motion of Met310, and approximately  $44 \text{ \AA}^3$  is due to side chain  
22 motion of Arg320, the combination of which increases the size of the polar pocket in the TR-  
23  $\beta$  LBD. This extra pliability also may explain the absence of ordered water in the polar  
24 pocket of TR- $\beta$  LBD bound to Triac or GC-1, which is in contrast to the ordered water  
25 found in the polar pocket of TR- $\alpha$  LBD bound to Dimit, IpBr2 or T3.  
26

Table 1\*

rTR- $\alpha$				
	<u>Dimit</u>	<u>Triac</u>	<u>IpBr2</u>	<u>T3</u>
TR LBD (volÅ <sup>3</sup> /areaÅ <sup>2</sup> )	590/456	589/440	601/474	605/472
Ligand (volÅ <sup>3</sup> /areaÅ <sup>2</sup> )	303/314	333/326	326/330	355/346
Complementarity	0.65	0.68	0.66	0.71

hTR- $\beta$		
	<u>GC-1</u>	<u>Triac</u>
TR LBD (volÅ <sup>3</sup> /areaÅ <sup>2</sup> )	665/575	625/474
Ligand (volÅ <sup>3</sup> /areaÅ <sup>2</sup> )	294/310	333/326
Complementarity	0.61	0.67

\*TR LBD volume and area are reported in Angstroms measured by GRASP. Complementarity is determined as defined in Lawrence *et al.*, *J. Mol. Biol.* 234:946-950 (1993).

Residue Ser277 in TR- $\alpha$  and the corresponding residue Asn331 of TR- $\beta$  also contribute to the volumetric differences observed in the polar pockets of these two TR isoforms. And substitution of the Asn331 of hTR- $\beta$  with serine has the affect of modifying ligand binding affinity of TR- $\beta$  so that it resembles that of TR- $\alpha$  (See Example 5). Taken together, differences in hydrogen bonding of atoms of the side chain group of Ser277 in TR- $\alpha$  and Asp331 in TR- $\beta$  extending from the equivalent backbone position in these TR LBDs and the more restricted polar pocket of the TR- $\alpha$  LBD further supports the concept of designing TR LBD isoform-specific ligands having substituents that fit spacially and preferentially into the polar pocket of either the TR- $\alpha$  or TR- $\beta$  LBDs. Exploitation of this difference provides an additional means for computational design of isoform-specific TR agonists and antagonists.

In terms of ligand design, these differences mean that for  $\beta$ -selective ligands, some or all of the following differences should be exploited:

1. The presence of a larger side chain asparagine.
2. The ability of the carbonyl group on the side chain to provide a strong hydrogen bond acceptor.
3. The ability of the amido group on the side chain to provide a two hydrogen bond donors.
4. Adjustment of polarity to reorganize the trapped water in the T3 pocket.

5. Greater size and flexibility of the polar pocket.

In terms of pharmaceutical design, these differences mean that for  $\alpha$ -selective ligands, some or all of the following differences should be exploited:

1. The presence of a smaller side group.
2. The ability of the hydroxyl on the  $-\text{CH}_2\text{OH}$  side group carbonyl group on the side chain to provide a weak hydrogen donor.
3. Adjustment of polarity to reorganize the trapped water in the T3 pocket.
4. Smaller size and limited flexibility of the polar pocket.

In both cases these differences can be exploited in a number of ways. For example, they can also be used with a software set for construction of novel organic molecules such as LUDI from Biosym-MSI. An example of designing TR- $\beta$  selective ligands is increasing the polarity of a ligand substituent located in the polar pocket of a TR LBD through addition of one or more ligand groups having a formal negative charge and/or negative dipole charge that interacts with a formal positive charge and/or positive dipole charge of a group in the polar pocket of the LBD. This exploits preferential interactions, such as with the additional positive charge contributed by Asn 331 in TR- $\beta$ . Another example of a TR- $\beta$  selective ligand is one that comprises one or more groups which fit spacially into the TR- $\beta$  LBD polar pocket. This exploits spacial differences between TR LBD isoforms, such as the larger and more flexible polar pocket of TR- $\beta$ .

## METHODS OF TREATMENT

The compounds of Formula 1 can be useful in medical treatments and exhibit biological activity which can be demonstrated in the following tests:

- (i) the induction of mitochondrial  $\alpha$ -glycerophosphate dehydrogenase (GPDH:EC 1.1.99.5). This assay is particularly useful since in certain species e.g. rats it is induced specifically by thyroid hormones and thyromimetics in a close-related manner in responsive tissues e.g. liver, kidney and the heart (Westerfield, W.W., Richert, D.A. and Ruegamer, W.R., *Endocrinology* (1965) 77:802). The assay allows direct measurement in rates of a thyroid hormone-like effect of compounds and in particular allows measurement of the direct thyroid hormone-like effect on the heart. Other measurements included parameters such as

1 heart rate and cardiac enzymes including  $\text{Ca}^{++}$  ATPase.  $\text{Na}^{++}/\text{K}^{+}$  ATPase, myosin isoforms  
2 and specific liver enzymes;

3 (ii) the elevation of basal metabolic rate as measured by the increase in whole  
4 body oxygen consumption (see e.g., Barker *et al.*, *Ann. N. Y. Acad. Sci.*, (1960) 86:545-  
5 562);

6 (iii) the stimulation of the rate of beating of atria isolated from animals previously  
7 dosed with thyromimetics (see e.g., Stephan *et al.*, *Biochem. Pharmacol.* (1992) 13:1969-  
8 1974; Yokoyama *et al.*, *J. Med. Chem.*, (1995) 38:695-707);

9 (iv) the change in total plasma cholesterol levels as determined using a cholesterol  
10 oxidase kit (for example, the Merck CHOD iodine colorimetric kit. see also, Stephan *et al.*  
11 (1992));

12 (v) the measurement of LDL (low density lipoprotein) and HDL (high density  
13 lipoprotein) cholesterol in lipoprotein fractions separated by ultracentrifugation; and p (vi) the  
14 change in total plasma triglyceride levels as determined using enzymatic color tests, for  
15 example the Merck System GPO-PAP method.

16 The compounds of Formula 1 can be found to exhibit selective thyromimetic activity  
17 in these tests,

18 (a) by increasing the metabolic rate of test animals, and raising hepatic GPDH  
19 levels at doses which do not significantly modify cardiac GPDH levels.

20 (b) by lowering plasma cholesterol and triglyceride levels, and the ratio of LDL to  
21 HDL cholesterol at doses which do not significantly modify cardiac GPDH levels.

22 The compounds of Formula 1 may therefore be used in therapy, in the treatment of  
23 conditions which can be alleviated by compounds which selectively mimic the effects of  
24 thyroid hormones in certain tissues whilst having little or no direct thyromimetic effect on the  
25 heart. For example, compounds of Formula 1 which raise hepatic GPDH levels and  
26 metabolic rate at doses which do not significantly modify cardiac GPDH levels are indicated  
27 in the treatment of obesity.

28 Agonists of Formula 1 will lower total plasma cholesterol, the ratio of LDL-  
29 cholesterol to HDL-cholesterol and triglyceride levels at doses which do not significantly  
30 modify cardiac GPDH levels are indicated for use as general antihyperlipidaemic  
31 (antihyperlipoproteinaemic) agents i.e. in the treatment of patients having elevated plasma



1 lipid (cholesterol and triglyceride) levels. In addition, in view of this effect on plasma  
2 cholesterol and triglyceride, they are also indicated for use as specific anti-  
3 hypercholesterolemic and anti-hypertriglyceridaemic agents.

4 Patients having elevated plasma lipid levels are considered at risk of developing  
5 coronary heart disease or other manifestations of atherosclerosis as a result of their high  
6 plasma cholesterol and/or triglyceride concentrations. Further, since LDL-cholesterol is  
7 believed to be the lipoprotein which induces atherosclerosis, and HDL-cholesterol believed to  
8 transport cholesterol from blood vessel walls to the liver and to prevent the build up of  
9 atherosclerotic plaque, anti-hyperlipidemic agents which lower the ratio of LDL-cholesterol  
10 to HDL cholesterol are indicated as anti-atherosclerotic agents, herein incorporated by  
11 reference U.S. patents 4,826,876 and 5,466,861.

12 The present invention also provides a method of producing selective thyromimetic  
13 activity in certain tissues except the heart which comprises administering to an animal in  
14 need thereof an effective amount to produce said activity of a compound of Formula 1 or a  
15 pharmaceutically acceptable salt thereof.

16 The present invention also relates to a method of lowering plasma lipid levels and a  
17 method of lowering the ratio of LDL-cholesterol to HDL-cholesterol levels by suitably  
18 administering a compound of this invention or a pharmaceutically acceptable salt thereof.

19 In addition, compounds of Formula 1 may be indicated in thyroid hormone  
20 replacement therapy in patients with compromised cardiac function.

21 In therapeutic use the compounds of the present invention are usually administered in  
22 a standard pharmaceutical composition.

23 The present invention therefore provides in a further aspect pharmaceutical  
24 compositions comprising a compound of Formula 1 or a pharmaceutically acceptable salt  
25 thereof and a pharmaceutically acceptable carrier. Such compositions include those suitable  
26 for oral, parenteral or rectal administration.

27

## 28 PHARMACEUTICAL COMPOSITIONS

29 Compounds of Formula 1 and their pharmaceutically acceptable salts which are active  
30 when given orally can be formulated as liquids for example syrups, suspensions or  
31 emulsions, tablets, capsules and lozenges.

1 A liquid composition will generally consist of a suspension or solution of the  
2 compound or pharmaceutically acceptable salt in a suitable liquid carrier(s), for example  
3 ethanol, glycerine, sorbitol, non-aqueous solvent such as polyethylene glycol, oils or water,  
4 with a suspending agent, preservative, surfactant, wetting agent, flavoring or coloring agent.  
5 Alternatively, a liquid formulation can be prepared from a reconstitutable powder.

6 For example a powder containing active compound, suspending agent, sucrose and a  
7 sweetener can be reconstituted with water to form a suspension; and a syrup can be prepared  
8 from a powder containing active ingredient, sucrose and a sweetener.

9 A composition in the form of a tablet can be prepared using any suitable  
10 pharmaceutical carrier(s) routinely used for preparing solid compositions. Examples of such  
11 carriers include magnesium stearate, starch, lactose, sucrose, microcrystalline cellulose and  
12 binders, for example polyvinylpyrrolidone. The tablet can also be provided with a color film  
13 coating, or color included as part of the carrier(s). In addition, active compound can be  
14 formulated in a controlled release dosage form as a tablet comprising a hydrophilic or  
15 hydrophobic matrix.

16 A composition in the form of a capsule can be prepared using routine encapsulation  
17 procedures, for example by incorporation of active compound and excipients into a hard  
18 gelatin capsule. Alternatively, a semi-solid matrix of active compound and high molecular  
19 weight polyethylene glycol can be prepared and filled into a hard gelatin capsule; or a  
20 solution of active compound in polyethylene glycol or a suspension in edible oil, for example  
21 liquid paraffin or fractionated coconut oil can be prepared and filled into a soft gelatin  
22 capsule. Compound of Formula 1 and their pharmaceutically acceptable salts which are  
23 active when given parenterally can be formulated for intramuscular or intravenous  
24 administration.

25 A typical composition for intra-muscular administration will consist of a suspension or  
26 solution of active ingredient in an oil, for example arachis oil or sesame oil. A typical  
27 composition for intravenous administration will consist of a sterile isotonic aqueous solution  
28 containing, for example active ingredient, dextrose, sodium chloride, a co-solvent, for  
29 example polyethylene glycol and, optionally, a chelating agent, for example ethylenediamine  
30 tetracetic acid and an anti-oxidant, for example, sodium metabisulphite. Alternatively, the

1 solution can be freeze dried and then reconstituted with a suitable solvent just prior to  
2 administration.

3 Compounds of structure (1) and their pharmaceutically acceptable salts which are  
4 active on rectal administration can be formulated as suppositories. A typical suppository  
5 formulation will generally consist of active ingredient with a binding and/or lubricating agent  
6 such as a gelatin or cocoa butter or other low melting vegetable or synthetic wax or fat.

7 Compounds of Formula 1 and their pharmaceutically acceptable salts which are active  
8 on topical administration can be formulated as transdermal compositions. Such compositions  
9 include, for example, a backing, active compound reservoir, a control membrane, liner and  
10 contact adhesive.

11 The typical daily dose of a compound of Formula 1 varies according to individual  
12 needs, the condition to be treated and with the route of administration. Suitable doses are in  
13 the general range of from 0.001 to 10 mg/kg bodyweight of the recipient per day.

14 Within this general dosage range, doses can be chosen at which the compounds of  
15 Formula 1 lower plasma cholesterol levels and raise metabolic rate with little or no direct  
16 effect on the heart. In general, but not exclusively, such doses will be in the range of from  
17 lower dose (0.001 to 0.5 mg/kg) to higher doses (0.5 to 10 mg/kg).

18 In addition, within the general dose range, doses can be chosen at which the  
19 compounds of Formula 1 lower plasma cholesterol levels and have little or no effect on the  
20 heart without raising metabolic rate. In general, but not exclusively, such doses will be in  
21 the range of from 0.001 to 0.5 mg/kg.

22 It is to be understood that the 2 sub ranges noted above are not mutually exclusive  
23 and that the particular activity encountered at a particular dose will depend on the nature of  
24 the compound of Formula 1 used.

25 Preferably, the compound of Formula 1 is in unit dosage form, for example, a tablet  
26 or a capsule so that the patient may self-administer a single dose. In general, unit doses  
27 contain in the range of from 0.05-100 mg of a compound of Formula 1. Preferred unit doses  
28 contain from 0.05 to 10 mg of a compound of Formula 1.

29 The active ingredient may be administered from 1 to 6 times a day. Thus daily doses  
30 are in general in the range of from 0.05 to 600 mg per day. Preferably, daily doses are in  
31 the range of from 0.05 to 100 mg per day. Most preferably from 0.05 to 5 mg per day.

## EXAMPLES

### EXAMPLE 1 - SYNTHESIS OF TR LIGANDS

Many TR ligands are known in the art, including T4 (thyroxine), T3, T2 and TS-9.

See Jorgensen, Thyroid Hormones and Analogs, in *β Hormonal Proteins and Peptides, Thyroid Hormones* 107-204 (Choh Hao Li ed., 1978), incorporated by reference herein.

The syntheses of several TR ligands are described below.

#### Synthesis of TS1, TS2, TS3, TS4, TS5

TS1, TS2, TS3, TS4 and TS5 and analogs thereof can all be prepared by simple acylation of the nitrogen atom of any thyronine analog, including T3 (3,5,3'-triiodo-L-thyronine), T4 (thyroxine) and 3,5-diiodothyronine. TS1 and TS2 are synthesized by reacting T3 with  $\text{Ph}_2\text{CHCO}_2\text{NHS}$  (N-hydroxy succinimide-2,2-diphenylacetate) and  $\text{C}_{16}\text{H}_{33}\text{CO}_2\text{NHS}$ , respectively. TS3 is synthesized by reacting T3 with FMOC-Cl (fluorenylmethyloxycarbonylchloride). TS4 is synthesized by reacting T3 with  $\text{tBOC}_2\text{O}$  (tBOC anhydride or di-t-butyldicarbonate). TS5, which differs from TS1-4 by having a -H instead of an -I at the  $\text{R}_3^1$  position, is synthesized by reacting 3,5-diiodothyronine with  $\text{tBOC}_2\text{O}$ . The general reaction scheme for TS1, TS2, TS3, TS4 and TS5 is depicted in FIG. 11. It should be noted that in the reaction scheme, both TS5 and its precursor both have a hydrogen rather than an iodine at the  $\text{R}_3^1$  position.

#### Synthesis of TS6 and TS7

TS6 is synthesized by reacting TS5 with paranitrophenylisocyanate. TS7 is synthesized by reacting TS6 with TFA (trifluoroacetic acid), which cleaves the tBOC group. These reactions are simple organic synthesis reactions that can be performed by anyone of ordinary skill in the art. The synthetic scheme for TS6 and TS7 is diagrammed in FIG. 12.

#### Synthesis of TS8

TS8 is synthesized by reacting TS5 with  $\text{Ph}_2\text{CHNH}_2$  (diphenylmethylamine) in the presence of triethylamine and any amide forming condensing reagent, such as TBTU (hydroxybenztriazoleuronium tetrafluoroborate) or HBTU (hydroxybenztriazoleuronium hexafluorophosphate). The synthesis scheme for TS8 is depicted in FIG. 13.

## 1 SYNTHESIS OF 3,5-DIIODO-3'-ISOPROPYLTHYRONINE DERIVATIVES

2 For designing a class of antagonists, it is important to have a hydrophobic group at  
3 the 3' position as well as an extension at the 5' position. Preferred hydrophobic groups at  
4 the 3' position include: methyl, benzyl, phenyl, iodo, and heterocyclic structures. The  
5 synthesis of a 3,5-diiodo-3'-isopropyl-5'-substituted thyronine is described below. The  
6 example provided describes the specific steps for synthesizing the TS10 compound, but this  
7 general reaction scheme can be used by one of ordinary skill in the art to synthesize any  
8 number of 3,5,-diiodo-3'-isopropyl-5'-substituted thyronine derivatives, which are  
9 characterized by having an extension at the 5' position. Additional compounds of this class  
10 can be synthesized using known organic synthesis techniques.

a 11 The synthesis of TS10 is described below and is depicted in <sup>FIGS. 14A-14B</sup> ~~FIG. 14~~. Numbers used  
12 in the reaction scheme for TS10 indicating the reaction product for each step are in  
13 parentheses.

14 2-Formyl-6-isopropylanisole (1): 2-formyl-6-isopropylanisole (10.0 g, 61 mmol), as-  
15 made by Casiraghi, *et al.* JCS Perkin I, 1862 (1980) (incorporated by reference), is added  
16 dropwise to a suspension of sodium hydride (3.7 g, 153 mmol) in 50 mL THF and 50 mL of  
17 DMF in a round bottom flask. The addition generates an exothermic reaction and formation  
18 of a gray solid. Methyl iodide (26.0 g, 183 mmol) is then added dropwise and the reaction  
19 mixture is stirred at room temperature for 5 hours. The reaction mixture is quenched with  
20 20 mL of water, then poured into 500 mL of water, and is extracted with ether (2 x 300  
21 mL). The ether layers are combined, washed with water (5 x 1000 mL), dried over  
22 magnesium sulfate and concentrated in vacuo to provide 10.2 g (94%) of the title compound,  
23 with the following <sup>1</sup>H NMR (CDCl<sub>3</sub>) properties: d 10.30 (s, 1H), 7.63 (d, 1H, J=3 Hz),  
24 7.50 (d, 1H, J=3 Hz), 7.13 (t, 1H, J=3 Hz), 3.81 (s, 3H), 3.31 (heptet, 1H, J=7.5 Hz),  
25 1.19 (d, 6H, J=7.5 Hz).

26  
27 2-(2-Hydroxynonyl)-6-isopropylanisole (not shown in scheme): Octylmagnesium  
28 chloride (8.4 mL, 16.9 mmol, 2.0 M) is added dropwise to a solution of 1 (1.5 g, 8.4 mmol)  
29 in 10 mL THF at -78°C. The reaction mixture is stirred for 2 hours with warming to room  
30 temperature. The reaction mixture is diluted with 50 mL ether and poured into 50 mL  
31 water. The ether layer is washed with brine (1 x 50 mL), dried over sodium sulfate, and

1 concentrated in vacuo. Flash chromatography (silica gel, 10% ether/hexane → 15%  
2 ether/hexane) provides 734 mg (30%) of the title compound with the following <sup>1</sup>H NMR  
3 (CDCl<sub>3</sub>) properties: δ 7.33-7.10 (m, 3H), 5.00 (br. s, 1H), 3.81 (s, 3H), 3.33 (heptet, 1H,  
4 J=7 Hz) 1.90-1.19 (m, 14H), 0.86 (t, 3H, J=6.5 Hz); HRMS (EI), found: 292.2404;  
5 calc'd: 292.2402.

6 2-nonyl-6-isopropylanisole (2): Compound 2 (663 mg, 2.3 mmol) is dissolved in  
7 solution of 5 mL ethanol and 5 mL acetic acid, and a spatula tip of palladium on carbon  
8 catalyst is added. The reaction mixture is then charged with hydrogen gas (using a simple  
9 balloon and needle) and the mixture is stirred at room temperature overnight. The next day,  
10 the reaction mixture is poured into ether (100 mL) and the ether layer is extracted with  
11 saturated sodium bicarbonate (3 x 100 mL). The ether layer is dried over sodium sulfate and  
12 concentrated *in vacuo* to provide 581 mg (91%) of (2) with the following <sup>1</sup>H NMR (CDCl<sub>3</sub>)  
13 properties: δ 7.14-7.00 (m, 3H), 3.75 (s, 3H), 3.36 (heptet, 1H, J=6.8 Hz), 2.63 (t, 2H,  
14 J=7.5 Hz), 1.68-1.15 (m, 14H), 0.86 (t, 3H, J=5.5 Hz); HRMS (EI), mass found:  
15 276.2459; calculated: 276.2453.

16 Thyronine adduct (4): Fuming nitric acid (0.071 mL) is added to 0.184 mL acetic  
17 anhydride chilled to -5°C. Iodine (66 mg) is added to this mixture followed by  
18 trifluoroacetic acid (0.124 mL). This mixture is stirred for 1 hour with warming to room  
19 temperature, at which point all of the iodine is dissolved. The reaction mixture was then  
20 concentrated *in vacuo* to provide an oily semi-solid material. The residue was dissolved in  
21 0.7 mL of acetic anhydride and cooled to -20°C. A solution of anisole (2) (581 mg, 2.1  
22 mmol) in 1.2 mL acetic anhydride and 0.58 mL TFA is added dropwise. The reaction  
23 mixture is stirred at -20° for 1 hour, then stirred overnight with warming to room  
24 temperature. The reaction mixture is partitioned between water and methylene chloride.  
25 The methylene chloride layer is dried over sodium sulfate and concentrated *in vacuo* to  
26 provide the iodonium salt (3) as an oil. This material is not purified or characterized, and is  
27 directly introduced into the coupling reaction.

28 N-Trifluoroacetyl-3,5-diiodotyrosine methyl ester (552 mg, 1.0 mmol) prepared  
29 according to the procedure of N. Lewis and P. Wallbank, *Synthesis* 1103 (1987)  
30 (incorporated by reference) and all of the crude iodonium salt (3) from above is dissolved in  
31 5 mL of anhydrous methanol. Diazabicyclo[5.4.0]undecane (DBU) (183 mg, 1.2 mmol) and

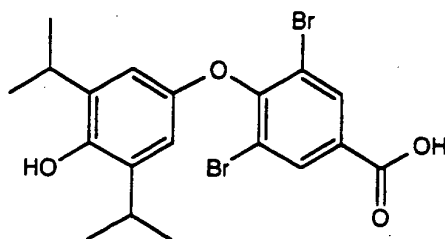
1 a spatula tip of copper-bronze are added and the resulting mixture is stirred at room  
2 temperature overnight. The next day, the reaction mixture is filtered, and the filtrate is  
3 concentrated *in vacuo*. The crude residue is purified by flash chromatography (silica gel,  
4 10% ethyl acetate/hexane) to provide 30 mg (4%) of the protected thyronine adduct (4).

5 Deprotected thyronine (TS10): The protected thyronine 4 (30 mg, 0.04 mmol) is  
6 dissolved in a mixture of 2.25 mL acetic acid and 2.25 mL 49% hydrobromic acid. The  
7 reaction mixture is heated to reflux for 5 hours. The reaction mixture is cooled to room  
8 temperature, and the solvents are removed *in vacuo*. Water is added to triturate the oily  
9 residue into a gray solid. This solid material is filtered, washed with water, and dried over  
10  $P_2O_5$  *in vacuo* to provide 24 mg (81%) of the title compound, TS10, with the following  $^1H$   
11 NMR ( $CDCl_3$ ) properties:  $\delta$  7.57 (s, 1H), 6.86 (s, 1H), 6.45 (s, 1H), 6.34 (s, 1H), 4.81  
12 (m, 1H), 3.86 (s, 3H), 3.71 (s, 3H), 3.33-3.05 (m, 3H), 2.58-2.47 (m, 2H), 1.62-0.76 (m,  
13 23H); MS (LSIMS):  $M^+ = 817.0$ .

14 As mentioned above, this reaction scheme can be modified by one of ordinary skill in  
15 the art to synthesize a class of compounds characterized by 3,5-diiodo-3'-isopropylthyronine  
16 derivatives, wherein (1) the 3' isopropyl group can be replaced with a hydrophobic group,  
17 including methyl, benzyl, phenyl, iodo, and heterocyclic structures, and (2) a wide variety of  
18 chemical structures can be incorporated at the 5' position, including alkyl groups, planar  
19 aryl, heterocyclic groups, or polar and/or charged groups.

20 The aldehyde (1) in the above reaction scheme is a versatile synthetic intermediate  
21 which allows for the attachment of a variety of chemical moieties to the 5' position of the  
22 final thyronine derivative. In addition, a variety of chemical reactions can be used to attach  
23 the chemical moieties. These reactions are well known in the art and include organometallic  
24 additions to the aldehyde (including Grignard reagents, organolithiums, etc.), reductive  
25 amination reactions of the aldehyde with a primary or secondary amine, and Wittig  
26 olefination reactions with a phosphorous ylid or stabilized phosphonate anion. Other  
27 possibilities include reduction of the aldehyde to a benzyl alcohol allowing for etherification  
28 reactions at the 5' position. As mentioned above, these methods allow for a wide variety of  
29 chemical structures to be incorporated at the 5' position of the final thyronine derivative,  
30 including alkyl groups, planar aryl, heterocyclic groups or polar and/or charged groups.

1    **Synthesis of 3, 5-dibromo-4-(3',5'-diisopropyl-4'-hydroxyphenoxy) benzoic acid**  
2    **(Compound 11).**



3  
4  
5  
6  
7  
8        (a) A mixture of 2,6-diisopropyl phenol (20 g, 0.11 mol), potassium carbonate (62 g,  
9    0.45 mol), acetone (160 ml) and methyl iodide (28 ml, 0.45 mole) is refluxed for three days.  
10    The reaction mixture is filtered through celite, evaporated, dissolved in ether, washed twice  
11    with 1M sodium hydroxide, dried over magnesium sulphate and concentrated to afford 15.1 g  
12    (0.08 mol, 70%) of 2,6-diisopropyl anisole as a slightly yellow oil.

13        (b) Fuming nitric acid (12.4 ml, 265 mmol) is added dropwise to 31.4 ml of acetic  
14    anhydride which is cooled in a dry ice/carbon tetrachloride bath. Iodine 11.3 g, 44.4 mmol)  
15    is added in one portion followed by dropwise addition of trifluoroacetic acid (20.5 ml, 266  
16    mmole). The reaction mixture is stirred at room temperature until all the iodine is dissolved.  
17    Nitrogen oxides are removed by flushing nitrogen into the vessel. The reaction mixture is  
18    concentrated, the residue is dissolved in 126 ml of acetic anhydride and is cooled in a dry  
19    ice/carbon tetrachloride bath. To the stirred solution 2,6-diisopropylanisole (51 g, 266  
20    mmol) in 150 ml of acetic anhydride and 22.6 ml of trifluoroacetic acid is added dropwise.  
21    The reaction mixture is left to stand at room temperature over night and then is concentrated.  
22    The residue is taken up in 150 ml of methanol and treated with 150 ml of 10% aqueous  
23    sodium bisulfite solution and 1 liter of 2M sodium borotetrafluoride solution. After the  
24    precipitate aggregates, petroleum ether is added and the supernatant is decanted. The  
25    precipitate is triturated with petroleum ether, filtered, washed with petroleum ether and dried  
26    at room temperature in vacuo. This affords 34 g (57 mmol, 65%) of bis(3,5-diisopropyl-4-  
27    methoxyphenyl)iodonium tetrafluoroborate as a white solid.

28        (c) To a stirred solution of 3,5-dibromo-4-hydroxybenzoic acid (12 g, 40.5 mmol) in  
29    250 ml of methanol, thionyl chloride (3 ml) is added dropwise. The reaction mixture is  
30    refluxed for five days, water is added and the precipitated product is filtered off. The  
31    residue is dissolved in ethyl acetate. From the aqueous phase, methanol is removed by



1 concentration. The aqueous phase is then saturated with sodium chloride, and extracted with  
2 ethyl acetate. The combined organic phases are dried over magnesium sulphate, filtered and  
3 concentrated. This gives 12.5 g (40.5 mmol, 100%) of 3,5-dibromo-4-hydroxymethyl  
4 benzoate as a white crystalline solid.

5 (d) The products obtained in steps b and c are reacted with each other according to  
6 the following protocol. To bis(3,5-diisopropyl-4-methoxyphenyl)iodonium tetrafluoroborate  
7 (2.86 g, 4.8 mmole) and copper bronze (0.42 g, 6.4 mmole) in 7 ml. of dichloromethane at  
8 0°C is added dropwise a solution of 3,5-dibromo-4-hydroxymethyl benzoate (1.0 g, 3.2  
9 mmole) and triethylamine (0.36 g, 3.5 mmole) in 5 ml of dichloromethane. The reaction  
10 mixture is stirred in the dark for eight days and then is filtered through celite. The filtrate is  
11 concentrated and the residue is purified by column chromatography (silica gel, 97:3  
12 petroleum ether/ethyl acetate) to give 0.62 g (1.2 mmole, 39%) of 3,5-dibromo-4-(3',5'-  
13 diisopropyl-4'-methoxyphenoxy)methyl benzoate as a solid.

14 (e) The product from step d (0.2 g, 0.4 mmole) is dissolved in 2 ml.  
15 dichloromethane, is put under nitrogen and is cooled at -40°C. To the stirred solution is  
16 added 1M BBr<sub>3</sub> (1.2 ml, 1.2 mmole) dropwise. The reaction mixture is allowed to reach  
17 room temperature and then is left over night. It is cooled to 0°C and then hydrolyzed with  
18 water. Dichloromethane is removed by concentration and the aqueous phase is extracted  
19 with ethyl acetate. The organic phase is washed with 1M hydrochloric acid and brine. Then  
20 it is dried over magnesium sulphate, filtered and concentrated. The residue is  
21 chromatographed (silica, 96:3.6:0.4 dichloromethane/methanol/acetic acid) producing 93 mg  
22 (0.2 mmole, 51 %) of 3,5-dibromo-4-(3',5'-diisopropyl-4'-hydroxyphenoxy)benzoic acid as a  
23 white solid. <sup>1</sup>H nmr (CDCl<sub>3</sub>) δ 1.23 (d, 12H, methyl), 3.11 (m, 2H, CH), 6.50 (s, 2H, 2,6-  
24 H) 8.33 (s, 2H, 2',6'-H).

25 Synthesis of addition ligands are described in U.S. Serial No. 08/877,792, filed June  
26 18, 1997 which is herein incorporated in its entirety by reference.

27 **TABLE 2** and **FIG. 15** depict the structures of several TR ligands in reference to  
28 Formula I.

TABLE 2

Cmpd	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	R <sub>1</sub>	R <sub>2</sub>	R <sub>6</sub>	R <sub>7</sub>
*T <sub>3</sub>	-I	-O-	-I	-I	-OH	-H	-CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H
*T <sub>4</sub>	-I	-O-	-I	-I	-OH	-I	-CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H
TS1	-I	-O-	-I	-I	-OH	-H	-CH <sub>2</sub> CH[NHCOCH <sub>2</sub> φ <sub>2</sub> ]CO <sub>2</sub> H
TS2	-I	-O-	-I	-I	-OH	-H	-CH <sub>2</sub> CH[NHCO(CH <sub>2</sub> ) <sub>15</sub> CH <sub>3</sub> ]CO <sub>2</sub> H
TS3	-I	-O-	-I	-I	-OH	-H	-CH <sub>2</sub> CH[NH-FMOC]CO <sub>2</sub> H
TS4	-I	-O-	-I	-I	-OH	-H	-CH <sub>2</sub> CH[NH-tBOC]CO <sub>2</sub> H
TS5	-I	-O-	-I	-H	-OH	-H	-CH <sub>2</sub> CH[NH-tBOC]CO <sub>2</sub> H
TS6	-I	-O-	-I	-H	-OC(O)NH=Ø <sub>p</sub> NO <sub>2</sub>	-H	-CH <sub>2</sub> CH[NH-tBOC]CO <sub>2</sub> H
TS7	-I	-O-	-I	-I	-OC(O)NH=NHØNO <sub>2</sub>	-H	-CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H
TS8	-I	-O-	-I	-H	-NH-CHØØ	-H	-CH <sub>2</sub> CH[NH-tBOC]CO <sub>2</sub> H
TS9	-I	-O-	-I	-IsoPr	-OH	-H	-CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H
TS10	-I	-O-	-I	-IsoPr	-OH	-(CH) <sub>8</sub> - CH <sub>3</sub>	-CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H

\* Prior Art Compound

-Ø: phenyl

-Ø<sub>p</sub>NO<sub>2</sub>: para nitro phenyl

## EXAMPLE 2 - RECEPTOR BINDING ASSAYS OF TR LIGANDS

To test the ability of synthesized TR ligands to bind to a thyroid receptor (TR), the binding affinity of a TR ligand for TR is assayed using TR's prepared from rat liver nuclei and <sup>125</sup>I T<sub>3</sub> as described in J.D. Apriletti, J.B. Baxter, and T.N. Lavin, *J. Biol. Chem.*, 263: 9409-9417 (1988). The apparent K<sub>d</sub>'s are calculated using the method described by Apriletti (1995) and Apriletti (1988). The apparent K<sub>d</sub>'s are presented in TABLE 3. The apparent K<sub>d</sub>'s (App.K<sub>d</sub>) are determined in the presence of the sample to be assayed, 1 nM [<sup>125</sup>I]T<sub>3</sub>, and 50μg/ml core histones, in buffer E (400 mM KCl, 200 mM potassium phosphate, pH 8.0, 0.5 mM EDTA, 1 mM MgCl<sub>2</sub>, 10% glycerol, 1 mM DTT) in a volume of 0.21 ml.

After incubation overnight at 4°C, 0.2 ml of the incubation mixture is loaded onto a Quick-Sep Sephadex G-25 column (2.7 x 0.9 cm, 1.7 ml bed volume) equilibrated with buffer E. The excluded peak of protein-bound [<sup>125</sup>I]T<sub>3</sub> is eluted with 1 ml of buffer E, collected in a test tube, and counted. Specific T<sub>3</sub> binding is calculated by subtracting nonspecific binding from total binding.

TABLE 3

Compound	App.Kd(nM)	Coactivation Assay RIP-140	EC <sub>50</sub> (M)
T <sub>3</sub>	0.06	+	10 <sup>-10</sup>
T <sub>4</sub>	2	+	10 <sup>-9</sup>
TS1	4	+	10 <sup>-7</sup>
TS2	1400	nd	nd
TS3	4	+	10 <sup>-8</sup>
TS4	8	+	nd
TS5	220	+	10 <sup>-6</sup>
TS6	> 10000	nd	nd
TS7	260	+	10 <sup>-7</sup>
TS8	6000	nd	nd
TS9	1	+	10 <sup>-10</sup>
TS10	400	+	10 <sup>-6</sup>

+: RIP-140 Binding

-: RIP-140 Binding

nd: Not Determined

### 1    **EXAMPLE 3 - INCREASED NUCLEAR PROTEIN COACTIVATION BY TR LIGANDS**

2            To test the ability of TR ligands to activate the binding of TR to the nuclear activation  
3 protein RIP-140 (a nuclear protein that can bind to nuclear receptors, such as the estrogen  
4 receptor), a TR ligand is liganded to TR and then incubated with RIP-140 as described in V.  
5 Cavailles, *et al.*, *EMBO J.*, 14(15):3741- 3751 (1995), which is incorporated by reference  
6 herein. In this assay, 35<sub>s</sub>-RIP-140 protein binds to liganded TR but not unliganded TR.  
7 Many TR 35<sub>s</sub> ligands can activate RIP-140 binding as shown in **TABLE 3**.

### 9    **EXAMPLE 4 - TR LIGAND BINDING AND TR ACTIVATION IN CULTURED CELLS**

10           To test TR activation of transcription in a cellular environment, TR ligands are  
11 assayed for their ability to activate a reporter gene, chloramphenicol transferase ("CAT"),  
12 which has a TR DNA binding sequence operatively linked to it. Either GC or L937 cells  
13 (available from the ATCC) can be used, respectively). In such assays, a TR ligand crosses  
14 the cell membrane, binds to the TR, and activates the TR, which in turn activates gene  
15 transcription of the CAT by binding the TR DNA binding region upstream of the CAT gene.  
16 The effective concentration for half maximal gene activation (EC<sub>50</sub>) is determined by assaying  
17 CAT gene activation at various concentrations as described herein and in the literature. The  
18 results of CAT gene activation experiments are shown in **TABLE 3**.

### 20    **CAT GENE ACTIVATION ASSAYS**

21           Functional response to thyroid hormone (3,5,3'-triiodo-L-thyronine, T<sub>3</sub>) and TR  
22 ligands is assessed either in a rat pituitary cell line, GC cells, that contain endogenous  
23 thyroid hormone receptors (TRs) or U937 cells that contain exogenous TRs expressed as  
24 known in the art. GC cells are grown in 10-cm dishes in RPMI 1640 with 10% newborn  
25 bovine serum, 2 mM glutamine, 50 units/ml penicillin and 50 µg/ml streptomycin. For  
26 transfections, cells are trypsinized, resuspended in buffer (PBS, 0.1% glucose) and mixed  
27 with a TREtkCAT plasmid (10 mg) or phage in 0.5 ml buffer (15±5 million cells) and  
28 electroporated using a Bio-Rad gene pulser at 0.33 kvolts and 960 mF. The TREtkCAT  
29 plasmid contains two copies of a T<sub>3</sub> response element (AGGTCAcaggAGGTCA) cloned in  
30 the Hind III site of the pUC19 polylinker immediately upstream of a minimal (-32/+45)  
31 thymidine kinase promoter linked to CAT (tkCAT) coding sequences. After electroporation,

1 cells are pooled in growth medium (RPMI with 10% charcoal-treated, hormone stripped.  
2 newborn bovine serum), plated in 6-well dishes and treated with either ethanol or hormone.  
3 CAT activity is determined 24 hours later as described D. C. Leitman, R. C. J. Ribeiro, E.  
4 R. Mackow, J. D. Baxter, B. L. West, *J. Biol. Chem.* 266, 9343 (1991), which is  
5 incorporated by reference herein.

6  
7 **EFFECT OF TS-10 ON THE TRANSCRIPTIONAL REGULATION OF THE DR4-ALP REPORTER**  
8 **GENE IN THE PRESENCE OR ABSENCE OF T3.**

9  
10 **Characteristics of the TRAF cells:** TRAFa1 are CHO K1 cells stably transformed with an  
11 expression vector encoding the human thyroid hormone receptor  $\alpha$  1 and a DR4,ALP  
12 reporter vector; TRAFb1 are CHO K1 cells stably transformed with an expression vector  
13 encoding the human thyroid hormone receptor  $\beta$ 1 and a DR4-ALP reporter vector.

14  
15 **Interpretation of the effect of compound TS-10 on the transcriptional regulation of the**  
16 **DR4-ALP reporter gene in the presence or absence of T3.**

17  
18 **TRAFa1 reporter cells:** TS-10 alone (open circles) induces a partial activation of the  
19 expression of the ALP reporter protein amounting to approximately 27% of the maximal  
20 effect by the natural thyroid hormone T3. In the presence of T3 (filled circles), TS-10 has a  
21 weak antagonistic effect. The EC50 concentration for the agonistic effect of TS-10 and the  
22 EC50 concentration for its T3 antagonistic effect, respectively, is indicated in FIG. 18.

23  
24 In FIG. 18, open and filled circles with dotted lines show the dose-dependent effect of TS-  
25 10/T3 on the toxicity marker (MTS/PMS), reduction of tetrazolium salt in the mitochondria,  
26 displayed on the right y-axis as optical density. There is no obvious toxic effect of TS-10 on  
27 the MTS-PMS marker but there is a clear effect on the morphology of the cells, as can be  
28 seen under the light microscope, at the highest concentration of TS-10 (32 mM) both in the  
29 absence and presence of T3, respectively (not shown in the figure).

1 **TRAFb1 reporter cells:** TS-10 alone (open circles) induces a partial activation of the  
2 expression of the ALP reporter protein amounting to approximately 35% of the maximal  
3 effect by T3. The EC50 concentration for the agonistic effect of TS-10 is indicated in FIG.  
4 19. In the presence of T3 (filled circles), TS-10 shows, if anything, a slight potentiation of  
5 the T3 effect on the expression of the ALP reporter protein. The T3 inhibitory effect of TS-  
6 10 at its highest concentration used (32 mM) is a toxic effect rather than T3 antagonism.

7  
8 In FIG. 19, open and filled circles with dotted lines show the dose-dependent effect of TS-  
9 10/T3 on the toxicity marker (MTS/PMS), reduction of tetrazolium salt in the mitochondria,  
10 displayed on the right y-axis as optical density. There is no obvious toxic effect of TS-10 on  
11 the MTS-PMS marker but a clear effect on the morphology of the cells can be observed,  
12 under the light microscope, at the highest concentration of TS-10 (32 mM) both in the  
13 absence and presence of T3, respectively (not shown in the figure).

14  
15 **HepG2 (HAF18) reporter cells:** TS-10 alone (open circles) induces a partial activation of  
16 the expression of the ALP reporter protein amounting to slightly more than 50% of the  
17 maximal effect by T3. The EC50 concentration for the agonistic effect of TS-10 is indicated  
18 in FIG. 20. In the presence of T3 (filled circles), TS-10 shows no effect i.e. no T3  
19 antagonism nor potentiation/additive effect to T3. Open and filled circles with dotted lines  
20 show the dose-dependent effect of TS-10/T3 on the toxicity marker (MTS/PMS), reduction  
21 of tetrazolium salt in the mitochondria, displayed on the right y-axis as optical density.  
22 There is no obvious toxic effect of TS-10 on the MTS/PMS marker or on the morphology of  
23 the cells, as can be observed using a light microscope, at any concentration of TS-10/T3  
24 used.

## 26 **Example 5 - Comparisons of Human TR- $\alpha$ and Human TR- $\beta$**

### 27 **Competition for [ $^{125}$ I]T<sub>3</sub> binding to TR LBD by T<sub>3</sub> and Triac**

28 The drug, Triac, is a thyroid hormone agonist. Triac is 3,5,3'-triiodothyroacetic acid  
29 and is described in Jorgensen, *Thyroid Hormones and Analogs in Hormonal Proteins and*  
30 *Peptides, Thyroid Hormones* at 150-151 (1978). Another compound that can be used in place  
31 of Triac is 3,5-diiodo-3'-isopropylthyroacetic acid. Competition assays are performed to

1 compare the displacement of [ $^{125}$ I]T<sub>3</sub> from binding with human TR- $\alpha$  LBD or human TR- $\beta$   
2 LBD by unlabeled T<sub>3</sub> or Triac. The results of such assays are depicted in FIG. 16.

3 Standard binding reactions are prepared containing 1 nM [ $^{125}$ I]T<sub>3</sub>, 30 fmol of human  
4 TR- $\alpha$  (empty symbols) or  $\beta$  (solid symbols), and various concentrations of competing  
5 unlabeled T<sub>3</sub> (circles) or Triac (triangles). Assays are performed in duplicate.

#### 6 7 Competition for [ $^{125}$ I]T<sub>3</sub> binding to variant TR LBD by T<sub>3</sub>, Triac and GC-1

8 The following assays residues involved in selective binding among TR isoforms.  
9 Competition assays are performed to compare the displacement of [ $^{125}$ I]T<sub>3</sub> from binding with  
10 wild-type human TR- $\alpha$  LBD or human TR- $\beta$  LBD, to a variant form of the TR LBDs by  
11 unlabeled T<sub>3</sub>, Triac or GC-1. A variant TR- $\alpha$  or TR- $\beta$  is constructed by substituting an  
12 amino acid found in the corresponding position of the other TR isoform. For example,  
13 asparagine 331 in human TR- $\beta$  corresponds to serine 277 in human TR- $\alpha$ . To test binding  
14 specificity contributed by this position, a variant human TR- $\beta$  is constructed that contains  
15 asparagine 331 substituted with a serine residue (designated Asn331Ser or N331S). Binding  
16 assays are described in *Apriletti et al.* (Protein Expression and Purification 6:363-370  
17 (1995)). The results of such assays are depicted in FIG. 27, and summarized in Table 4  
18 below.

19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
TABLE 4  
Effect of TR- $\beta$  Substitution N331S on Binding Affinity

Ligand	Native TR- $\alpha$	Native TR- $\beta$	Mutant TR- $\beta$
T3	20 pM	60 pM	100 pM
T4	600	3000	ND
Triac	20	20	100
IpBr <sub>2</sub>	17	ND	ND
Dimit	6000	8000	ND
GC-1	200	40	400

1 Competition curves comparing wildtype TR- $\beta$  versus the variant TR- $\beta$  N331S for binding  
2 T3, Triac or GC-1 show that the affinity of the mutant receptor for Triac was reduced to  
3 approximately the same as for T3 (vs. 3-fold greater in wild type) so that the relative  
4 affinities are similar to wild-type TR- $\alpha$ . The affinity for GC-1 was also reduced to several  
5 fold less than T3, as is seen with TR- $\alpha$ .

6 Comparison of the affinity of TR- $\beta$  variant N331S to the native TRs for selected  
7 ligands is as follows:

8 Native TR- $\alpha$  for various ligands (T3, T4, Triac, IpBr2, Dimit, GC-1):

9 IpBr<sub>2</sub> > Triac = T3 > GC-1 > T4 > Dimit

10 Native TR- $\beta$  (T3, T4, Triac, Dimit, GC-1)

11 Triac > GC-1  $\geq$  T3 > T4 > Dimit

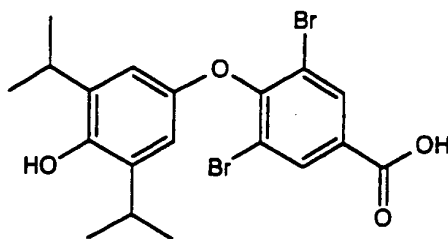
12 Variant TR- $\beta$  (N331S) (T3, Triac, GC-1)

13 Triac = T3 > GC-1.

#### 14 Scatchard Analysis of [<sup>125</sup>I]T<sub>3</sub> Binding to TR

15 Human TR- $\alpha$  (left panel) or human TR- $\beta$  (right panel) is assayed for T<sub>3</sub> binding in the  
16 presence of increasing concentrations of [<sup>125</sup>I]T<sub>3</sub>. The apparent equilibrium dissociation  
17 constant (20 pM for  $\alpha$  and 67 pM for  $\beta$ ) is calculated by linear regression analysis and is  
18 depicted in FIG. 17. *FIGS. 17A-17B*

19 *a* 20 3, 5-DIBROMO-4-(3',5'-DIISOPROPYL-4'-HYDROXYPHENOXY) BENZOIC ACID IS A  
21 TR- $\alpha$  SELECTIVE SYNTHETIC LIGAND.



22  
23  
24  
25  
26  
27  
28  
29 3, 5-dibromo-4-(3',5'-diisopropyl-4'-hydroxyphenoxy) benzoic acid (Compound 11),  
30 the structure of which is drawn above, is assayed for binding to the two different isoforms of  
31 the TR, TR- $\alpha$  and TR- $\beta$ . Compound 11 exhibits an IC<sub>50</sub> of 1.6  $\mu$ M for binding to TR- $\alpha$



1 and an IC<sub>50</sub> of 0.91  $\mu$ M for binding to TR- $\beta$ . Assays for determining selective binding to  
2 the TR- $\alpha$  or TR- $\beta$  LBD can include reporter assays, as described herein. See also  
3 Hollenberg, *et al.*, *J. Biol. Chem.*, (1995) 270(24):14274-14280.

#### 5 **EXAMPLE 6 - PREPARATION AND PURIFICATION OF A TR- $\alpha$ LBD**

6 Rat TR- $\alpha$  LBD, residues Met122 - Val410, is purified from *E. coli* ("LBD-122/410").  
7 The expression vector encoding the rat TR- $\alpha$  LBD is freshly transfected into *E. coli* strain  
8 BL21(DE3) and grown at 22°C in a 50-liter fermenter using 2x LB medium. At an A<sub>600</sub> of  
9 2.5-3, IPTG is added to 0.5 mM and growth is continued for 3 h before harvesting. The  
10 bacterial pellet is quickly frozen in liquid nitrogen and stored at -70°C until processed.  
11 Extraction and purification steps are carried out at 4°C. The bacteria are thawed in  
12 extraction buffer (20MM Hepes, pH 8.-, 1 mM EDTA, 0.1% MTG, 0.1 mM PMSF, and  
13 10% glycerol) at a ratio of 10 ml buffer/g bacteria. Bacteria are lysed by incubation for 15  
14 min. with 0.2 mg/ml lysozyme and sonicated at maximum power while simultaneously  
15 homogenized with a Brinkmann homogenizer (Model PT 10/35 with generator PTA 35/2)  
16 until the solution loses its viscosity. After centrifugation for 10 min at 10,000 g, the  
17 supernatant is adjusted to 0.4 M KCl, treated with 0.6% PEI to precipitate fragmented DNA,  
18 and centrifuged for 10 min at 10,000 g. The rat TR- $\alpha$  LBD in the supernatant is then  
19 precipitated with 50% ammonium sulfate and centrifuged for 10 min at 10,000 g. The  
20 precipitate is resuspended with buffer B (20 mM Hepes, pH 8.0, 1 mM EDTA, 1 mM DTT,  
21 0.1 mM PMSF, 0.01% Lubrol, and 10% glycerol) to a final conductivity of 9 mS/cm  
22 (approx. 0.7 M ammonium sulfate) and centrifuged 1 h at 100,000g. The supernatant is  
23 frozen in liquid nitrogen and stored at -70°C.

24 The crude extract is thawed, bound with a tracer amount of [<sup>125</sup>I]T<sub>3</sub>, and loaded  
25 directly onto a phenyl-Toyopearl hydrophobic interaction column (2.6 x 18 cm, 95 ml bed  
26 volume) at 1.5 ml/min. The column is eluted with a 2-h gradient from 0.7 ammonium  
27 sulfate, no glycerol to no salt, 20% glycerol in buffer C (20 mM Hepes, pH 8.0, 0.5 mM  
28 EDTA, 1 mM DTT, 0.2 mM PMSF). The rat TR- $\alpha$  LBD prebound to tracer [<sup>125</sup>I]T<sub>3</sub> (less  
29 than 0.005% of total rat TR- $\alpha$  LBD) is detected using a flow-through gamma emission  
30 detector, whereas unliganded rat TR- $\alpha$  LBD is assayed by postcolumn [<sup>125</sup>I]T<sub>3</sub> binding assays  
31 (described herein).

1 The phenyl-Toyopearl unliganded rat TR- $\alpha$  LBD peak fractions are pooled, diluted  
2 with buffer B to a conductivity of 0.5 mS/cm (equivalent to approx. 20 mM ammonium  
3 sulfate), loaded onto a TSK-DEAE anion-exchange column (2 x 15 cm, 47 ml bed volume) at  
4 4 ml/min, and eluted with a 60-min gradient from 50 to 200 mM NaCl in buffer B.

5 The unliganded rat TR- $\alpha$  LBD peak fractions from TSK-DEAE are pooled, diluted  
6 twofold with buffer B, loaded at 0.75 ml/min on a TSK-heparin HPLC column (0.8 x 7.5  
7 cm, 3 ml bed volume), and eluted with a 50 to 400 mM NaCl gradient in buffer B.

8 The pool of unliganded rat TR- $\alpha$  LBD peak fractions from the TSK-heparin column is  
9 adjusted to 0.7 M ammonium sulfate, loaded at 0.75 ml/min on a TSK-phenyl HPLC column  
10 (0.8 x 7.5 cm, 3 ml bed volume), and eluted with a 60-min gradient from 0.7 M ammonium  
11 sulfate without glycerol to no salt with 20% glycerol in buffer C. The fractions containing  
12 unliganded rat TR- $\alpha$  LBD are pooled and incubated with a five fold excess of hormone for 1  
13 h, the salt concentration is adjusted to 0.7 M ammonium sulfate, and the sample is reloaded  
14 and chromatographed on the same column as described above.

#### 16 EXAMPLE 7 - CRYSTALLIZATION OF LIGANDED TR- $\alpha$ LBD

17 Material from a single LBD-122/410 preparation is divided into batches, and  
18 quantitatively bound with one of the following ligands: Dimit, T<sub>3</sub>, or Triac IpBr<sub>2</sub>  
19 (3,5dibromo-3'-isopropylthyronine) for the final purification step.

20 To maintain full saturation of rat TR- $\alpha$  LBD with a ligand, and to prepare the  
21 complex for crystallization, the ligand-bound rat TR- $\alpha$  LBD is concentrated and desalted in  
22 an Amicon Centricon-10 microconcentrator (McGrath et al, *Biotechniques*, (1989) 7:246-247,  
23 incorporated by reference herein), using 10 mM Hepes (pH 7.0), 3.0 mM DTT, and 1.0 nM  
24 to 10 nM ligand.

25 Factorial crystallization screening trials (Jancarik & Kim, *J. Appl. Crystallogr.* (1991)  
26 24:409-411, incorporated by reference herein) are carried out for rat TR- $\alpha$  LBD bound to  
27 selected ligands using hanging-drop vapor diffusion at 17°C (with 1  $\mu$ l protein solution, 1  $\mu$ l  
28 precipitant solution and a 0.5 ml reservoir using silanized coverslip: (McPherson, *Preparation*  
29 *and Analysis of Protein Crystals* (1982), incorporated by reference herein). Rat TR- $\alpha$  LBD  
30 is not stable at 4°C and is stored at -80°C, where it maintains its avidity for hormone and its  
31 crystallizability for approximately two to three months. These procedures are carried out as

described in McGrath, M.E. *et al.*, *J. Mol. Biol.* (1994) 237:236-239 (incorporated by reference). Crystals are obtained in condition 21 of the screening trials (Jancarik & Kim 1991) and conditions are then optimized. Wedge-shaped crystals are reproducibly obtained with hanging-drop vapor fusion at 22°C with 15% 2-methyl-2,4-pentanediol (MPD), 0.2 M ammonium acetate and 0.1 M sodium cacodylate (pH 6.7), 3 mM DTT, with 2 µl protein solution, 1 µl precipitant solution and a 0.6 ml reservoir using silanized coverslip, and with 8.7 mg/ml (Dimit), 5.5 mg/ml (IpBr<sub>2</sub>), 5 mg/ml (Triac), or 2.3 mg/ml (T<sub>3</sub>) over a period of three days. Under these conditions, diffraction quality crystals (dimension 0.5 x 0.2 x 0.0075 mm<sup>3</sup>) can be grown at ambient temperature (22°C). The best crystals have a limiting dimension of approximately 100 µm and are obtained at a protein concentration between 2.3 and 8.7 mg/ml in the presence of 3 mM DTT. The crystals are of the monoclinic space group C2, with one monomer in the asymmetric unit.

#### EXAMPLE 8 - CRYSTALLIZATION OF HUMAN TR-β LBD COMPLEXED WITH T3, TRIAC, OR GC-1

Human TR-β LBD complexed with T<sub>3</sub>, Triac, or GC-1 are purified according to the same procedures described above for the rat TR-α LBD, with the following modifications.

The expression of human TR-β LBD differs from the rat TR-α LBD in that the human TR-β LBD residues extend from the amino acid at position 716 through the amino acid at position 1022, according to the amino acid numbering scheme for the various nuclear receptor LBDs depicted in ~~FIG. 3~~. <sup>FIGS. 3A-3E</sup> ~~FIG. 3~~ illustrates a numbering scheme applicable to all of the nuclear receptors listed as well as to any additional homologous nuclear receptors.

The vertical lines on ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup> at position 725 and at position 1025 delineate the preferred minimum amino acid sequence necessary to obtain adequate binding of ligand. The amino acid sequence from position 716 to position 1022 according to the numbering scheme of ~~FIG. 3~~ <sup>FIGS. 3A-3E</sup> corresponds to the amino acid positions 202 to 461 according to the conventional numbering of the amino acid sequence of human TR-β which is publicly available. Also, the human TR-β LBD is expressed with a histidine tag, as described in Crowe *et al.*, *Methods in Molecular Biology* (1994) 31:371-387, incorporated by reference herein.

The purification of human TR-β LBD is the same as that described above for the rat TR-α LBD with the following exceptions. First, before the purification step using the

hydrophobic interaction column, a step is added in which the expressed human TR- $\beta$  LBD is purified using a nickel NTA column (commercially available from Qiagen, Chatsworth, CA) according to manufacturer's instructions, and eluted with 200 mM imidazole. The second difference is that in the purification of the human TR- $\beta$  LBD, the purification step using a heparin column is omitted.

The crystallization of human TR- $\beta$  LBD bound to T<sub>3</sub>, Triac or GC-1 is as follows. Crystals are obtained in condition 7 of the factorial screen using hanging drops as before at ambient temperature (22°C) using the factorial crystallization screening trials of Jancarik & Kim (1991) and using the commercially available product from Hampton Research, Riverside). The following are optimum conditions: hexagonal bipyrimidal crystals are grown at 4°C for 2-3 days from hanging drops containing 1.0-1.2 M sodium acetate (pH unadjusted) and 0.1 M sodium cacodylate (pH 7.4), 3 mM DTT, with either a 1  $\mu$ l protein solution, 1  $\mu$ l precipitant solution or 2  $\mu$ l protein solution, 1  $\mu$ l precipitant solution and a 0.6 ml reservoir using silanized coverslip, at a protein concentration of 7-10 mg/ml. The best crystals have a limiting dimension of 200  $\mu$ m. The following are optimum conditions for crystallization of the TR- $\beta$  LBD with GC-1: hexagonal bipyrimidal crystals are grown at 4°C for 2-3 days from hanging drops containing 0.8-1.0M sodium acetate (pH unadjusted), 50-200nM sodium succinate, and 0.1M sodium cacodylate (pH 7.2), 3mM DTT, 1  $\mu$ l protein solution, 1  $\mu$ l precipitant solution and a 0.6ml reservoir using silanized coverslip, at a protein concentration of 7-10 mg/ml. The best crystals have a limiting dimension of 200 $\mu$ M. The unit cell dimensions are cell length a=b=68.73, cell length c=130.09. The unit cell angles are  $\alpha=90^\circ$ ,  $\beta=90^\circ$ ,  $\gamma=120^\circ$ .

The crystal system for human TR- $\beta$  LBD bound to T<sub>3</sub>, Triac or GC-1 is trigonal with the space group p3<sub>1</sub>21. The unit cell dimensions are cell length a = cell length b = 68.448 angstroms, cell length c = 130.559 angstroms. The angles are  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$ , gamma = 120°.

#### EXAMPLE 9 - DETERMINATION OF LIGANDED TR- $\alpha$ LBD AND TR- $\beta$ CRYSTAL STRUCTURES

Data from each cocrystal (Rat TR- $\alpha$  LBD with Dimit, T3 and IpBr2; Human TR- $\beta$  LBD with Triac and GC-1) is measured on a Mar area detector at Stanford Synchrotron

1 Radiation Laboratory beamline 7-1 ( $\lambda = 1.08$  angstroms) using 1.2° oscillations. Data from  
2 the cocrystal of the hTR- $\beta$  LBD with Triac is measured on a Mar area detector at Stanford  
3 Synchrotron Radiations Laboratory beamline 7-1 ( $\lambda = 1.08$  angstroms) using 1.0  
4 oscillations. Data from the cocrystal of the hTR- $\beta$  LBD with GC-1 is measured on a R-axis  
5 II area detector on a Rigaku rotating Cu anode (50kV, 300mA). The crystals are transferred  
6 into a cryosolvent containing 1.2M sodium acetate, 0.1M sodium cacodylate, and 15%  
7 glycerol followed by a second transfer into 30% glycerol, then flash frozen in liquid  
8 nitrogen. An orientation matrix for each crystal is obtained using DENZO. The reflections  
9 are integrated with DENZO (commercially available from Molecular Structure Corp., The  
10 Woodlands, Texas) and are scaled with SCALEPACK (as described in Otwinowski, Z,  
11 *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing,"* 56-62 (SERC  
12 Daresbury Laboratory, Warrington, UK 1993) incorporated by reference).

13 For rTR- $\alpha$  cocrystals, data from the  $T_3$  cocrystal is measured with the  $b^*$  axis  
14 approximately parallel with the spindle. The crystals are flash frozen at -178°C in a nitrogen  
15 gas stream with the MPD mother liquor serving as the cryosolvent. An orientation matrix  
16 for each crystal is determined using REFI X (Kabsch, W., *J. Appl. Crystallogr.* (1993)  
17 26:795-800 incorporated by reference). Reflections are integrated with DENZO, and are  
18 scaled with SCALEPACK.

19 For the  $T_3$  data set, Bijvoet pairs are kept separate, and are locally scaled using  
20 MADSYS (W. Hendrickson (Columbia University) and W. Weis (Stanford University)).

21 Cocrystals prepared from the three isosteric ligands are isomorphous. MIR analysis is  
22 performed using programs from the CCP4 suite (Collaborative Computational Project, N.R.  
23 *Acta Crystallogr.* (1994) D50:760-763, incorporated by reference herein). Difference  
24 Pattersons are calculated for both  $T_3$  and IpBr<sub>2</sub>, taking the Dimit cocrystal as the parent. The  
25 positions of the three iodine atoms in the  $T_3$  difference Patterson are unambiguously  
26 determined from the Harker section of the density map as peaks of  $11\sigma$  above background.  
27 The positions for the two bromine atoms in the IpBr<sub>2</sub> cocrystals, are located independently,  
28 as peaks  $8\sigma$  above the noise level. Phases for the LBD-122/410 are calculated from the  
29 solution to the IpBr<sub>2</sub> difference Patterson, and are used to confirm the location of the unique  
30 third iodine of the  $T_3$  cocrystal. Halogen positions are refined with MLPHARE, including the  
31 anomalous contributions from the iodine atoms (Otwinowski, Z, *Proceedings of the CCPR*

1 *Study Weekend 80-86* (SERC Daresbury Laboratory, Warrington, UK 1991)). The MIRAS  
 2 phases are improved through solvent flattening/histogram matching using DM (Cowtan, K..  
 3 *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* (1994) 31: 34-38,  
 4 incorporated by reference herein).

5 A model of the LBD-122/410 with Dimit bound is built with the program O from the  
 6 solvent flattened MIRAS 2.5 angstrom electron density map (Jones *et al.*, *Acta Crystallogr.*  
 7 (1991) A 47:110-119, incorporated by reference herein). The initial model, without ligand,  
 8 (R<sub>cryst</sub> = 40.1%), is refined using least-squares protocols with XPLOR. The Dimit ligand  
 9 is built into unambiguous Fo-Fc difference density during the following round. Subsequent  
 10 refinement employs both least-squares and simulated annealing protocols with XPLOR  
 11 (Brunger *et al.*, *Science* (1987) 235:458-460), incorporated by reference herein). Individual  
 12 atomic B-factors are refined isotropically. As defined in PROCHECK, all residues are in  
 13 allowed main-chain torsion angle regions as described in Laskowski *et al.*, *J. Appl.*  
 14 *Crystallogr.*, (1993) 26:283-291, incorporated by reference herein. The current model is  
 15 missing 34 residues (Met<sub>122</sub>-Gln<sub>156</sub>) at the N-terminus, and 5 residues (Glu<sub>406</sub>-Val<sub>410</sub>) at the C-  
 16 terminus.

17 In addition, the following residues are not modeled beyond C $\beta$  due to poor density:  
 18 184, 186, 190, 198, 206, 209, 240, 301, 330, 337, 340, 343, 359, and 395. The average B-  
 19 value for protein atoms is 34.5 Å<sup>2</sup>. The final model consists of the LBD-122/410, residues  
 20 Arg<sub>157</sub>-Ser<sub>183</sub>, Trp<sub>185</sub>-Gly<sub>197</sub>, Ser<sub>199</sub>-Asp<sub>206</sub> and Asp<sub>208</sub>-Phe<sub>405</sub>; three cacodylate-modified  
 21 cysteines: Cys<sub>334</sub>, Cys<sub>380</sub> and Cys<sub>392</sub>; and 73 solvent molecules modeled as water (2003  
 22 atoms).

$$23 \quad *R_{\text{sym}} = 100 \times \sum_{hkl} \sum_i |I_i - I| / \sum_{hkl} \sum_i I_i$$

$$24 \quad \dagger R_{\text{der}} = 100 \times \sum_{hkl} |F_{\text{PH}} - F_{\text{H}}| / \sum_{hkl} |F_{\text{P}}|$$

25 The occupancy for the two bromine sites is set to 35 electrons. The occupancies of the iodine  
 26 sites are relative to this value.

27 §Phasing power =  $\langle \text{FH} \rangle / \langle \epsilon \rangle$ , where  $\langle \text{FH} \rangle$  is the mean calculated heavy atom structure factor  
 28 amplitude and  $\langle \epsilon \rangle$  is the mean estimated lack of closure.

29 ||R<sub>culis</sub> =  $\langle \epsilon \rangle / \langle \text{iso} \rangle$ , where  $\langle \epsilon \rangle$  is the mean estimated lack of closure and  $\langle \text{iso} \rangle$  is the  
 30 isomorphous difference.

1  $R_{\text{cryst}} = 100 \times \sum_{hkl} |F_o - F_c| / \sum_{hkl} |F_o|$  where  $F_o$  and  $F_c$  are the observed and calculated  
 2 structure factor amplitudes (for data  $F/\sigma > 2$ ). The  $R_{\text{free}}$  was calculated using 3% of the  
 3 data, chosen randomly, and omitted from the refinement.

4  $\text{\% Correlation coefficient} = \sum_{hkl} (|F_o| - |F_c|) \times (|F_o| - |F_c|) / \sum_{hkl} (|F_o| - |F_c|)^2 \times \sum_{hkl}$   
 5  $(|F_o| - |F_c|)^2$

#### 7 **EXAMPLE 10. PHASING OF THE rTR- $\alpha$ LBD AND hTR- $\beta$ LBD COMPLEX WITH TRIAC**

8 Due to the possible non-isomorphism of the rTR $\alpha$  LBD complex with Triac, a  
 9 molecular replacement solution is determined using AMORE (Navaza, J., *Acta*  
 10 *Crystallographica Section A-Fundamentals of Crystallography* (1994) 50:157-63 from a  
 11 starting model consisting of rTR $\alpha$  LBD complex with T3, but with the ligand, all water  
 12 molecules, and the following residues omitted: Asn 179, Arg228, Arg262, Arg266, and Ser  
 13 277. Strong peaks are obtained in both the rotation and translation searches, with no  
 14 significant ( $> 0.5$  times the top peak) false solutions observed (Table 6). Strong positive  
 15 density present in both the anomalous and conventional difference Fourier maps confirm the  
 16 solution. Maps are calculated using sigma-A weighted coefficients output by REFMAC  
 17 (Murshudov, *et al.* "Application of Maximum Likelihood Refinements," in *Refinement of*  
 18 *Protein Structures, Proceedings of Daresbury Study Weekend* (1996)) after 15 cycles of  
 19 maximum likelihood refinement. Triac, the omitted residues, and water molecules 503, 504,  
 20 534 (following the numbering convention for the TR complex with T3) are built into the  
 21 resulting difference density using O (Jones *et. al.*); the conformations of these residues are  
 22 further confirmed in a simulated-annealing omit map (Brunger *et. al.*). The complete model  
 23 is then refined using positional least-squares, simulated annealing, and restrained, grouped B  
 24 factor refinement in XPLOR to an  $R_{\text{cryst}}$  of 23.6% and an  $R_{\text{free}}$  of 24.1%

25 Phasing of a related LBD using the structure of the rTR- $\alpha$  LBD is conducted as  
 26 follows. A molecular replacement solution for the hTR- $\beta$  LBD complex with Triac is  
 27 determined using AMORE from a starting model consisting of the rTR- $\alpha$  LBD complexed  
 28 with T3, but with the ligand and all water molecules omitted. Strong peaks are obtained in  
 29 both the rotation and translation searches, with no significant ( $> 0.5$  times the top peak) false  
 30 solutions (Table 7). Strong positive density present in both the anomalous and conventional  
 31 difference Fourier maps confirm the solution. Initial maps are calculated using sigma-A

1 weighted coefficients output by REFMAC after 9 cycles of maximum likelihood refinement.  
2 The real-space fit for each residues was calculated using OOPS (Kleywegt, GJ and Jones.  
3 TA, OOPS-a-daisy, ESF/CCP4 Newsletter 30, June 1994, pp. 20-24) and the residues with a  
4 real-space fit less than 2 standard deviations below the mean removed: Ala253-Lys263;  
5 Glu245-Leu250. To reduce bias, the following residues were modeled as alanine: Arg282,  
6 Arg316, Arg 320, Asn 331. Cycles of rebuilding and positional least-squares, simulated  
7 annealing, and restrained, grouped B factor refinement with XPLOR produce a model with  
8 an  $R_{\text{cryst}}$  of 25.3 and an  $R_{\text{free}}$  of 28.9%. The final model consists of hTR- $\beta$  LBD residues  
9 Glu202-Gln252, Val264-Glu460; three cacodylate-modified cysteines with the cacodylate  
10 moiety modeled as free arsenic: Cys294, Cys298, Cys388, and Cys434; and 35 solvent  
11 molecules modeled as water.

12

#### 13 **EXAMPLE 11. CONNECTING QSAR WITH STRUCTURE IN THE THYROID HORMONE** 14 **RECEPTOR**

15 The conclusions of classic thyroid hormone receptor quantitative structure-activity  
16 relationships may be summarized as follows:

- 17 1) the  $R_4'$ -hydroxyl group functions as a hydrogen bond donor;
- 18 2) the amino-propionic acid interacts electrostatically through the carboxylate  
19 anion with a positively charged residue from the receptor;
- 20 3) the preferences of  $R_3/R_5$  substituent are  $I > Br > Me > > H$ ;
- 21 4) the preferences of the  $R_3'$ -substituent are  $Ipr > I > Br > Me > > H$ .

22 The structure of the thyroid hormone receptor ligand binding domain complexed with the  
23 agonists T3, IpBr<sub>2</sub>, Dimit, Triac, and GC1 as provided herein, permits:

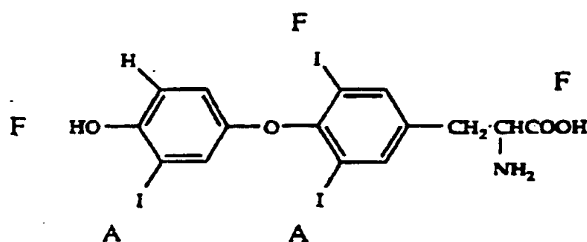
- 24 1) the identification of receptor determinants of binding at the level of the  
25 hydrogen bond;
- 26 2) the association of these determinants with the predictions of classic thyroid  
27 hormone receptor QSAR; and
- 28 3) prediction as to which determinants of binding are rigid, and which are  
29 flexible, for both the ligand and the receptor.

30 This classification for the agonists of the type ( $R_1$ =amino-propionic, acetic acid;  
31  $R_3, R_5=I, Br, Me$ ;  $R_3'=Ipr, I$ ) is given below (for the representative ligand T<sub>3</sub>);



1 F = Fiducial (always satisfied)

2 A = Adjustable



12 Based upon the methods and data described herein, the following is an embodiment of  
13 the computational methods of the invention, which permit design of nuclear receptor ligands  
14 based upon interactions between the structure of the amino acid residues of the receptor LBD  
15 and the four different ligands described herein. The small molecule structures for the ligands  
16 can be obtained from Cambridge Structural Database (CSD), and three dimensional models  
17 can be constructed using the methods described throughout the specification. The following  
18 are factors to consider in designing synthetic ligands:

19 1) Histidine 381 acts as a hydrogen bond acceptor for the  $R_4'$  hydroxyl, with the  
20 optimal tautomer maintained by water molecules. See FIG. 23 and FIG 24. Histidine is the  
21 only hydrophilic residue in this hydrophobic pocket that surrounds the  $R_4'$  substituent.  
22 Histidine can be either a hydrogen bond acceptor or donor, depending on its tautomeric state.  
23 It is preferably a hydrogen bond donor, but can tolerate being a hydrogen bond acceptor, as  
24 for example, when there is a methoxy at the  $R_4'$  position of the ligand;

25 2) Arginines 228, 262, and 266 interact directly and through water-mediated  
26 hydrogen bonds with the  $R_1$ -substituent, with the electrostatic interaction provided by  
27 Arginine 266 (as in the Triac complex). This polar pocket is illustrated by FIG. 23 - FIG.  
28 25. FIG. 23 depicts  $T_3$  in the  $TR\alpha$  ligand binding cavity, where  $T_3$ 's amino-propionic  $R_1$ -  
29 substituent interacts with Arg 228, HOH502, H9H503 and HOH504 via hydrogen bonds.  
30 FIG. 24 depicts Triac in the ligand binding cavity, with its -COOH  $R_1$  substituent in the  
31 polar pocket. In FIG. 24, Arg 228 no longer shares a hydrogen bond with the ligand, but  
the -COOH  $R_1$  substituent forms hydrogen bonds with Arg 266. FIG. 25 superimposes  $T_3$

1 and Triac in the ligand binding cavity and shows several positionally unchanged amino acids  
2 and water molecules, and selected changed interacting amino acids and water molecules.  
3 The three figures illustrate parts of the polar pocket that can change and those parts that do  
4 not move upon binding of different ligands. For example, the Arg 262 at the top of the  
5 polar pocket does not move, even when the  $R_1$  substituent has changed from a -COOH to an  
6 aminopropionic acid group. However, the other two Arginines, Arg 228 and Arg 266,  
7 demonstrate flexibility in the polar pocket to respond to the change in the size or chemical  
8 nature of the  $R_1$  substituent.

9 3) Inner and outer pockets for the  $R_3/R_5$  substituents are formed by Ser260,  
10 Ala263, Ile299; and Phe 218, Ile221, Ile222, respectively. See FIGS. 21 and 22. The  
11 inner pocket is filled by either the  $R_3$  or the  $R_5$  substituent, regardless of the size of the  
12 substituent, and may act as a binding determinant by positioning the ligand in the receptor.  
13 Optimally, the inner pocket amino acids interact with an  $R_3$  or  $R_5$  substituent that is no  
14 larger than an iodo group. If the inner pocket is filled by the  $R_3$  substituent, then the outer  
15 pocket interacts with the  $R_5$  substituent and vice versa. The outer pocket can adjust to the  
16 size of its substituent through main chain motion centered at the break in helix 3 (Lys220-  
17 Ile221), suggesting that the bending of H3, and motion of the N-terminal portion of H3, may  
18 represent a conformational change induced on ligand binding. The outer pocket has greater  
19 flexibility than does the inner pocket in terms of accommodating a larger substituent group.

20 4) A pocket for the  $R_3'$ -substituent is formed by Phe 215, Gly290, Met388. The  
21 pocket is incompletely filled by the  $R_3'$ -iodo substituent, and accommodates the slightly  
22 larger 3'-isopropyl substituent by movement of the flexible Met388 side chain and the H7/H8  
23 loop. This pocket can accommodate  $R_3'$  substituents that are even larger than isopropyl, for  
24 example, a phenyl group.

25 The above information will facilitate the design of high affinity agonists and  
26 antagonists by improving automated QSAR methodologies and informing manual modeling of  
27 pharmaceutical lead compounds. For example, the inclusion of discrete water molecules  
28 provides a complete description of hydrogen bonding in the polar pocket for use with  
29 pharmacophore development; also, the identification of mobile and immobile residues within  
30 the receptor suggests physically reasonable constraints for use in molecular  
31 mechanics/dynamics calculations.

## 1 EXAMPLE 12. DESIGN OF AN INCREASED AFFINITY LIGAND

2 Direct interaction between the receptor and the ligand is limited in the polar pocket,  
3 which interacts with the  $R_1$  substituent. While the lack of complementarity may contain  
4 implications for biological regulation, it also provides an opportunity for increasing affinity  
5 by optimizing the interaction between the amino acids of the polar pocket and the  $R_1$   
6 substituent of a synthetic ligand. The structure of the receptor-ligand interactions described  
7 herein enables design of an increased affinity synthetic ligand having two complementary  
8 modifications:

9 1) Remove the positively charged amine. The strongly positive electrostatic  
10 potential predicted for the polar pocket suggests that the positively charged amine of the  
11 aminopropionic acid  $R_1$  substituent may be detrimental to binding. Suitable groups for  
12 substitution are suggested by the nature of nearby hydrogen bond partners: for example, Thr  
13 275 O or Ser 277 N. See e.g. Tables in Appendix 2. For example, any any negatively  
14 charged substituent would be compatible for interacting with the amino acids of the polar  
15 pocket, including carboxylates, carbonyl, phosphonates, and sulfates, comprising 0 to 4  
16 carbons. Another example of an  $R_1$  substitution is an oxamic acid that replaces the amine of  
17 the naturally occurring ligand with one or more carbonyl groups.

18 2) Incorporate hydrogen bond acceptor and donor groups into the  $R_1$ -substituent  
19 to provide broader interactions with the polar pocket scaffold. Such hydrogen bond acceptor  
20 and donor groups incorporated into the  $R_1$ -substituent will allow interactions that would  
21 otherwise occur with water molecules in the polar pocket. Specific waters include HOH 504  
22 (hydrogen bonds with Ala 225 O and Arg 262 NH); and HOH 503 hydrogen bonds with Asn  
23 179 OD1, Ala 180 N), both of which are present in all four complexes (TR LBD complexed  
24 with T3, TR LBD complexed with IpBr<sub>2</sub>, TR LBD complexed with Dimit and TR LBD  
25 complexed with Triac). Analysis of the hydrogen bonding network in the polar pocket  
26 suggests replacement of HOH 504 with a hydrogen bond acceptor, and HOH 503 with an  
27 hydrogen bond donor (although the chemical nature of asparagine probably permits flexibility  
28 at this site). Thus, incorporating a hydrogen bond acceptor in an  $R_1$  substituent that could  
29 take the place of the HOH504 or incorporating a hydrogen bond acceptor in an  $R_1$   
30 substituent that could positionally replace the HOH503, or a combination thereof, are  
31 methods of designing novel synthetic TR ligands.

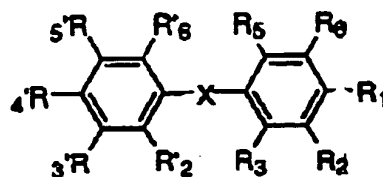
1        These two design approaches can be used separately or in combination to design  
2 synthetic ligands, including those in Table 5 (below).

3        A corollary to this approach is to design specific interactions to the residues Arg262  
4 and Asn 179. The goal is to build in interactions to these residues by designing ligands that  
5 have  $R_1$  substituents that form hydrogen bonds with water molecules or charged residues in  
6 the polar pocket.

7        High-affinity ligands also may be designed and selected using small molecules that  
8 bind to proximal subsites of the target nuclear hormone receptor that are identified in a  
9 structure-based screen and then linked together in their experimentally determined bound  
10 orientations. Such a method has been described in design of high-affinity ligands for the  
11 FK506 binding protein (FKBP), stromelysin, gelatinase A, and human papillomavirus E2  
12 (Hajduk *et al.*, *Science* 278:497-499 (1997)), which reference and its references are  
13 incorporated herein by reference. The preferred small molecules for screening are  
14 compounds of Formula I or derivatives thereof. For example, a compound of Formula I ( $\phi$ -  
15  $X-\phi$ ) or a derivative thereof ( $\phi-X$  or  $X-\phi$ ) is screened for binding a target nuclear hormone  
16 receptor LBD. Proximal subsites of the nuclear hormone receptor include the hydrophobic  
17 and polar pockets of the LBD, and subsites extended therefrom. As an example, Fourier  
18 transformation or nuclear magnetic resonance (NMR) -based structure screens can be used.  
19 When a NMR-based screen is used, binding can be detected from the amide chemical shift  
20 changes observed in two-dimensional heteronuclear single quantum correlation (HSQC)  
21 spectra acquired in the presence and absence of added compound. Once two ligands are  
22 identified that bind to the receptor, the crystal or solution structure of the ternary complex is  
23 determined. From the structural information, a compound is synthesized which links the two  
24 ligands, where the linker is selected based on structural information. The new compound is  
25 then screened for binding affinity, for example, using a binding assay as described herein.  
26 Only a few linked ligands need to be synthesized and screened when using this approach.

27        Compounds of the invention also may be iteratively designed from structural  
28 information of the compounds described above using other structure-based design/modeling  
29 techniques (Jackson, R.C., *Contributions of protein structure-based drug design to cancer*  
30 *chemotherapy. Seminars in Oncology*, 1997, 24(2):L164-172; and Jones, T.R., *et al.*, *J.*  
31 *Med. Chem.*, 1996 39(4):904-917).

Table 5: Synthetic TR Ligands



R1	R2	R3	R5	R6	X	R'2	R'3	R'4	R'5	R'6
CO <sub>2</sub> H	H	Me	Me	H	O	H	Me	OH	Me	H
CH <sub>2</sub> CO <sub>2</sub> H		I	I		S		Et	SH	Et	
CH <sub>2</sub> CH <sub>2</sub> CO <sub>2</sub> H		Br	Br				nPr	NH <sub>2</sub>	nPr	
CH <sub>2</sub> CH(NH <sub>2</sub> )CO <sub>2</sub> H		Cl	Cl				iPr		iPr	
OCH <sub>2</sub> CO <sub>2</sub> H		Et	Et				Ph		nBu	
OCH <sub>2</sub> CH <sub>2</sub> CO <sub>2</sub> H		OH	OH				I		nPen	
NHCH <sub>2</sub> CO <sub>2</sub> H		NH <sub>2</sub>	NH <sub>2</sub>				Br		nHex	
NHCH <sub>2</sub> CH <sub>2</sub> CO <sub>2</sub> H		SH	SH				Cl		Ph	
CH <sub>2</sub> COCOCOC <sub>2</sub> H									hetero	
									cycle	
NHCOCOCOC <sub>2</sub> H									aryl	
COCOC <sub>2</sub> H										
CF <sub>2</sub> CO <sub>2</sub> H										
COCH <sub>2</sub> CO <sub>2</sub> H										

Any combination of the above substituents in the biphenyl ether scaffold structure shown above may result in a potentially pharmacologically useful ligand for the thyroid hormone receptor. These novel ligands may be antagonists of the thyroid receptor.

TABLE 6: TR- $\alpha$  LBD-122/410

	Dimit	T3	IpBr <sub>2</sub>	Triac
1				
2				
3	<b>Data collection</b>			
4	<b>Cell dimensions</b>			
5	a (Å)	117.16	117.19	117.18
6	b (Å)	80.52	80.20	80.12
7	c (Å)	63.21	63.23	63.13
8	$\beta$ (°)	120.58	120.60	120.69
9	Resolution (Å)	2.2	2.0	2.1
10	Obs. Reflections, (no.)	57031	64424	66877
11	Unique Reflections, (no.)	22327	21023	23966
12				18453
13	Completeness, (%)	87.0	82.4	93.7
14	*R <sub>sym</sub> (%)	3.9	3.5	4.5
15	<b>Phasing (15.0 - 2.5Å)</b>			
16	†R <sub>der</sub> (%)	-	19.6	11.6
17	No. of sites	-	3	2
18	‡Occupancy	-	44.6 (19.8)	35.0
19	(Anomalous)	-	50.2 (23.7)	35.0
20			39.2 (22.3)	
21	§F <sub>H</sub> /E			
22	centric (acentric)			
23	15.0-5.0 Å	-	3.67 (4.61)	2.25 (3.09)
24	5.0-3.0 Å	-	2.23 (2.75)	1.25 (1.85)
25	3.0-2.5 Å	-	1.64 (1.99)	1.15 (1.57)
26	R <sub>Cullis</sub> (%)			
27	15.0-5.0 Å	-	33	44
28	5.0-3.0 Å	-	45	63
29	3.0-2.5 Å	-	60	65
30	Mean figure of merit	0.62	-	-
31	MR Phasing			
32	(10-3.5Å)			
33	Rotation Search:			$\Theta_1 = 309.37$
34	Euler Angles (°)			$\Theta_2 = 48.96$
35				$\Theta_3 = 127.28$
36	§ correlation coefficient			34.3
37	Translation Search:			x = 0.1571
38	Fractional coordinates			
39				y = 0.000

1					$z = 0.3421$
2	§ correlation				65.8
3	Coefficient				
4	'R factor				31.2
5	Refinement	15.0-2.2	5.0 - 2.0	15.0 - 2.2	25-2.5
6	Resolution (Å)				
7	$R_{\text{cryst}}$ (%)	20.5	22.1	21.4	23.6
8	$R_{\text{free}}$ (%)	22.7	24.0	22.4	24.1

TABLE 7: TR- $\beta$  LBD-202/461

12		Triac	T3	GC1
13	Data collection			
14	Space Group		P3121	
15	Cell dimensions			
16	a (Å)	68.9	68.45	68.73
17	c (Å)	131.5	130.56	130.09
18	Resolution (Å)	2.4	3.1	2.8
19	Obs. Reflections, (no.)	80196	55103	54104
20	Unique Reflections, (no.)	14277	6847	8987
21	Coverage (%)	97.0	95.7	97.1
22	* $R_{\text{sym}}$ (%)	5.1	4.6	5.5
23	MR Phasing (15.0 - 3.5Å)			
24	Rotation Search	$\Theta_1 = 39.13$		
25	Euler Angles (°)	$\Theta_2 = 68.00$		
26		$\Theta_3 = 323.6$		
27	§ correlation coefficient	21.6		
28	(Highest false peak)	(10.8)		
29	Translation Search	$x = 0.748$		
30	Fractional Coordinates	$y = 0.158$		
31		$z = 0.167$		
32	§ correlation coefficient	57.5		
33	(Highest false peak)	(38.7)		
34		0.612		
35	*R factor	40.7	40.8	
36	Refinement			
37	Resolution (Å)	30-2.4		30-2.9
38	$R_{\text{cryst}}$ (%)	25.3		27.3
39	$R_{\text{free}}$ (%)	28.9		33.4

1  
2 All publications and patent applications mentioned in this specification are herein  
3 incorporated by reference to the same extent as if each individual publication or patent  
4 application was specifically and individually indicated to be incorporated by reference. The  
5 nuclear receptor ligands, particularly the TR ligands, of these references are herein  
6 incorporated by reference and can be optionally excluded from the claimed compounds with a  
7 proviso.

8 Headings and subheadings are presented only for the convenience of the reader and  
9 should not be used to construe the meaning of terms used within such headings and  
10 subheadings.

11 The invention now being fully described, it will be apparent to one of ordinary skill -  
12 in the art that many changes and modifications can be made thereto without departing from  
13 the spirit or scope of the appended claims.